



UNIVERSITY OF IOANNINA
SCHOOL OF ECONOMICS AND BUSINESS
DEPARTMENT OF ECONOMICS

**Female labour-force participation
and socioeconomic factors, with emphasis on health:
Dynamic models estimated with UK data**

by
Anastasia Sofia Grigoriadou

**A dissertation submitted in partial fulfilment of the requirements for the degree of
DOCTOR OF PHILOSOPHY**

Ioannina 2021

Υποψήφια Διδάκτορας: Γρηγοριάδου Αναστασία Σοφία,

Τμήμα Οικονομικών Επιστημών, Πανεπιστήμιο Ιωαννίνων.

Θέμα Διατριβής: Female labour-force participation and socioeconomic factors, with emphasis on health: Dynamic models estimated with UK data

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ

Επιβλέπων

Χατζηνικολάου Δημήτριος, Αναπληρωτής Καθηγητής, Τμήμα Οικονομικών Επιστημών, Πανεπιστήμιο Ιωαννίνων.

Μέλη

Βλάσσης Μηνάς, Καθηγητής, Μέλος της Τριμελούς επιτροπής

Παπαπανάγος Χαράλαμπος, Καθηγητής, Μέλος της Τριμελούς επιτροπής

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Βλάσσης Μηνάς, Καθηγητής, Τμήμα Οικονομικών Επιστημών, Πανεπιστήμιο Κρήτης, Μέλος της Τριμελούς Επιτροπής

Γιαννακόπουλος Νικόλαος, Αναπληρωτής Καθηγητής, Τμήμα Οικονομικών Επιστημών, Πανεπιστήμιο Πατρών

Γναρδέλλης Χαράλαμπος, Καθηγητής, Τμήμα Ζωικής Παραγωγής, Αλειάς & Υδατοκαλλιεργειών, Πανεπιστήμιο Πατρών

Μπένος Νικόλαος, Αναπληρωτής Καθηγητής, Τμήμα Οικονομικών Επιστημών, Πανεπιστήμιο Ιωαννίνων

Μπίλιας Ιωάννης, Καθηγητής, Τμήμα Διεθνών και Ευρωπαϊκών Οικονομικών Σπουδών, Οικονομικό Πανεπιστήμιο Αθηνών

Παπαπανάγος Χαράλαμπος, Καθηγητής, Τμήμα Βαλκανικών, Σλαβικών και Ανατολικών Σπουδών, Πανεπιστήμιο Μακεδονίας, Μέλος της Τριμελούς επιτροπής

Χατζηνικολάου Δημήτριος, Αναπληρωτής Καθηγητής, Επιβλέπων, Τμήμα Οικονομικών Επιστημών, Πανεπιστήμιο Ιωαννίνων

Acknowledgements

Special thanks go to my major supervisor, Associate Professor Dimitris Hatzinikolaou, for giving me the opportunity to undertake the PhD program. I am grateful for the meetings, comments, and continuous support throughout this academic journey and also for giving me the chance and joy to do auxiliary teaching. I also thank Mark Brown, my postgraduate professor at the University of Manchester, for strengthening my idea of continuing my education and pursuing a PhD. Besides, I wish to thank my PhD thesis committee for comments and suggestions. My heartfelt gratitude to Vasilis Gialamas, professor at the National and Kapodistrian University of Athens, for his essential guidance. Last, but not least, I would like to thank my son, Stefanos, for inspiring me and motivating me in his own unique way.

Dedication

To her

Abstract

Using a balanced panel of 331 married women aged 39-45 from the British Household Survey for the years 1991-2002, we seek to investigate the dynamics of female labour supply with emphasis on how health affects the labour force participation decisions of women. State dependence in participation status, unobserved individual heterogeneity, autocorrelation in the error term and the initial conditions are taken into account through the adaptation of appropriate dynamic binary discrete choice models. The framework has been described by Heckman (1981b), Hyslop (1999), Keane and Sauer (2009) and Bartolucci and Nigro (2010). These four approaches are compared with a dynamic pooled probit model, a dynamic random effects probit model, and a static random effects probit model to assess the robustness of the results. We also implemented Mundlak's (1978) correlated random effects approximation, embedded in Hyslop's (1999) approach.

To the best of our knowledge, our two contributions to the literature are as follows: (1) it is the first study that examines the impact of both physical and mental health on women's work decisions as we use several aspects of health as proxies for health (in particular, self-assessed health, whether health limits daily activities and health problems related to anxiety and depression), and (2) it is the first study that examines female labour supply within a framework determined by Bartolucci and Nigro (2010).

In all approaches, the coefficient estimate on lagged participation status reflects positive state dependence and is statistically significant. Hyslop's (1999) model exhibits the largest estimate of state dependence and indicates that ignoring autocorrelation in the error term would lead to underestimation of state dependence. The estimated average partial effect of lagged participation by Hyslop's (1999) approach indicates that, averaged across all women and all time periods, the probability of a woman participating in the current year is 60.37 per cent higher if the woman was participating in the previous year than if she was not.

The significant effect of health-related variables is identified in all estimators. The findings suggest that women have a greater probability of participation when they do not

report limiting health problems and when they claim excellent, very good, good or fair health, compared to women who claim poor or very poor health. In addition, Mundlak's (1978) correlated random effects approximation, embedded in Hyslop's (1999) approach, indicated that limiting health problems are correlated with individual characteristics that reduce the probability of participation. Moreover, the predicted probabilities demonstrate the strong effect of the presence of young children and of health problems on the probability of participating. The presence of at least one young child in the household diminishes the probability of a woman's participation. The younger the children are, the higher the effect is.

In addition, an additional year of work experience and an additional year of education increase significantly the probability of participation. Moreover, coefficient estimates for permanent non-labour income are also statistically significant and have the expected sign. Finally, we find a negative added worker effect, verifying similar findings from empirical studies in the UK.

**Συμμετοχή των γυναικών στην αγορά εργασίας και
προσδιοριστικοί παράγοντες, με έμφαση στην υγεία. Δυναμικά μοντέλα
εκτιμημένα με δεδομένα από το Ηνωμένο Βασίλειο.**

ΠΕΡΙΛΗΨΗ

Χρησιμοποιώντας ένα ισορροπημένο panel 331 παντρεμένων γυναικών ηλικίας 39-45 ετών από τη British Household Survey για τα έτη 1991-2002, επιδιώκουμε να διερευνήσουμε τη δυναμική της προσφοράς εργασίας γυναικών με έμφαση στο πώς η υγεία επηρεάζει τις εργασιακές αποφάσεις των γυναικών. Η state dependence (η εξάρτηση του τρέχοντος στάτους συμμετοχής στο εργατικό δυναμικό από αυτό της προηγούμενης περιόδου) στη συμμετοχή των γυναικών στην αγορά εργασίας, η μη παρατηρούμενη ατομική ετερογένεια, η αυτοσυσχέτιση του όρου σφάλματος και οι αρχικές συνθήκες λαμβάνονται υπόψη μέσω της προσαρμογής κατάλληλων δυναμικών μοντέλων διακριτών επιλογών. Το πλαίσιο περιγράφεται από τους Heckman (1981b), Hyslop (1999), Keane and Sauer (2009) και Bartolucci and Nigro (2010). Αυτές οι τέσσερις προσεγγίσεις συγκρίνονται με ένα δυναμικό pooled probit, ένα δυναμικό μοντέλο τυχαίων επιδράσεων probit και ένα στατικό μοντέλο τυχαίων επιδράσεων probit για την αξιολόγηση της ευρωστίας των αποτελεσμάτων. Εφαρμόσαμε επίσης την προσέγγιση των συσχετισμένων τυχαίων επιδράσεων του Mundlak (1978), ενσωματώνοντάς τη στη προσέγγιση του Hyslop (1999).

Από όσο γνωρίζουμε, η διατριβή συνεισφέρει στη βιβλιογραφία ως εξής: (1) είναι η πρώτη μελέτη που εξετάζει τον αντίκτυπο όχι μόνο της σωματικής, αλλά και της ψυχικής υγείας στην προσφορά εργασίας των γυναικών, καθώς χρησιμοποιούμε διάφορες πτυχές της υγείας στη θέση της μεταβλητής «υγεία» (ως proxies, δηλαδή, συγκεκριμένα, η αυτο-αξιολογούμενη υγεία, το κατά πόσον η υγεία περιορίζει τις καθημερινές δραστηριότητες, καθώς και προβλήματα υγείας που σχετίζονται με το άγχος και την κατάθλιψη), και (2)

είναι η πρώτη μελέτη που εξετάζει την προσφορά εργασίας των γυναικών στα πλαίσια του υποδείγματος των Bartolucci και Nigro (2010).

Σε όλες τις προσεγγίσεις, η εκτίμηση του συντελεστή για την κατάσταση συμμετοχής στην αγορά εργασίας με χρονική υστέρηση μίας περιόδου αντανakλά θετική state dependence και είναι στατιστικά σημαντική. Το μοντέλο του Hyslop (1999) παρουσιάζει τη μεγαλύτερη εκτίμηση της state dependence στη συμμετοχή στην αγορά εργασίας και υποδεικνύει ότι η παράβλεψη της αυτοσυσχέτισης στον όρο σφάλματος θα οδηγούσε σε υποεκτίμηση της state dependence για τη συμμετοχή στην αγορά εργασίας. Η εκτιμώμενη μέση μερική επίδραση της συμμετοχής στην αγορά εργασίας με χρονική υστέρηση μίας περιόδου από την προσέγγιση του Hyslop (1999) δείχνει ότι, κατά μέσο όρο σε όλες τις γυναίκες και σε όλες τις χρονικές περιόδους, η πιθανότητα μιας γυναίκας να συμμετέχει στην αγορά εργασίας το τρέχον έτος είναι 60,37% υψηλότερη εάν η γυναίκα συμμετείχε στην αγορά εργασίας το προηγούμενο έτος από ό,τι αν δεν συμμετείχε.

Η σημαντική επίδραση των μεταβλητών που σχετίζονται με την υγεία εντοπίζεται σε όλους τους εκτιμητές. Τα ευρήματα δείχνουν ότι οι γυναίκες έχουν μεγαλύτερη πιθανότητα συμμετοχής στην αγορά εργασίας όταν δεν αναφέρουν περιοριστικά προβλήματα υγείας και όταν δηλώνουν άριστη, πολύ καλή, καλή ή μέτρια υγεία, σε σύγκριση με τις γυναίκες που δηλώνουν ότι έχουν κακή ή πολύ κακή υγεία. Επιπλέον, η προσέγγιση του Mundlak (1978), ενσωματωμένη στη προσέγγιση του Hyslop (1999), έδειξε ότι τα περιοριστικά προβλήματα υγείας συσχετίζονται με ατομικά χαρακτηριστικά που μειώνουν την πιθανότητα συμμετοχής στην αγορά εργασίας. Επιπλέον, οι προβλεπόμενες πιθανότητες καταδεικνύουν την ισχυρή επίδραση της παρουσίας μικρών παιδιών και των προβλημάτων υγείας στην πιθανότητα συμμετοχής στην αγορά εργασίας. Η παρουσία τουλάχιστον ενός μικρού παιδιού στο νοικοκυριό μειώνει την πιθανότητα συμμετοχής μιας γυναίκας στην αγορά εργασίας. Όσο μικρότερα είναι τα παιδιά, τόσο μεγαλύτερη είναι η επίδραση.

Ακόμη, ένα επιπλέον έτος εργασιακής εμπειρίας και ένα επιπλέον έτος εκπαίδευσης αυξάνουν σημαντικά την πιθανότητα συμμετοχής στην αγορά εργασίας. Επιπλέον, οι εκτιμήσεις των συντελεστών για το μόνιμο εναλλακτικό εισόδημα, δηλαδή

το εισόδημα που προέρχεται από πηγές πλην της εργασίας της γυναίκας, είναι επίσης στατιστικά σημαντικές και έχουν το αναμενόμενο πρόσημο. Τέλος, βρίσκουμε μία αρνητική επίδραση του επιπρόσθετου εργάτη, επαληθεύοντας παρόμοια ευρήματα από εμπειρικές μελέτες στο Ηνωμένο Βασίλειο.

Contents

Acknowledgements.....	ii
Dedication.....	iii
Abstract	iv
Contents.....	x
List of tables	XIII
List of figures.....	XIV
Chapter 1 Introduction	1
Chapter 2 Review Of Literature.....	14
2.1 Introduction.....	14
2.2 Figures And Trends.....	16
2.3 Women’s work decisions	21
2.3.1 Reduced-form approximations in female labour supply studies.....	21
2.3.2 Structural approximations in female labour supply studies	29
2.4 Health.....	31
2.4.1 Health. Definitions and measures.	31
2.4.2 Endogeneity of health in labour supply models.....	33
2.4.3 Links between labour supply and health.....	39
2.4.4 UK studies on the links between labour supply and health.....	45

Chapter 3 Theoretical Framework And Econometric Approaches	51
3.1 Introduction	51
3.2 A theoretical framework for intertemporal labour supply decision	52
3.3 Econometric approaches	58
3.3.1 Random-Effects Approaches	60
3.3.2 Fixed-Effects Approach	64
3.3.3 Mundlak's Approach	70
3.4 Average partial effects (APEs)	74
 Chapter 4 The Data.....	 73
4.1 Introduction	73
4.2 The British Household Panel Survey	73
4.3 Data preparation and variables of the dataset	75
4.3.1 Data preparation	75
4.3.2 Variables of the dataset	77
4.4 Summary Statistics of the dataset	85
4.5 Summary/Overview	95
 Chapter 5 Estimation Results	 97
5.1 Introduction	97
5.2 Results and Discussion	98
5.3 Summary	124

Chapter 6 Conclusion.....	126
References.....	131
Appendices.....	140
Appendix A.1.....	140
Appendix A.1.1	140
Appendix A.1.2	142
Appendix A.1.3	144
Appendix A.1.4	145
Appendix A.1.5	145
Appendix A.1.6	146
Appendix A.1.7	147
Appendix A.1.8	148
Appendix A.1.9	148
Appendix A.2.....	153
Appendix A.3.....	153
Appendix A.4.....	154
Appendix A.5.....	155
Appendix A.6.....	159
Appendix A.7.....	160

List of Tables

Table 4.1: Current economic activity (wave 1).....	78
Table 4.2 Descriptive statistics: wave 1 and wave 12	86
Table 4.3 Descriptive statistics: all waves.....	87
Table 4.4 Employment state transitions	88
Table 4.5 Employment state by self-assessed health	89
Table 4.6 Employment state by health limitations	89
Table 4.7 Employment states by wave.....	91
Table 4.8 Health status by wave	91
Table 4.9 Health limitations by wave.....	92
Table 4.10 Anxiety/depression by wave.....	92
Table 5.1 Static and dynamic probit models	100
Table 5.2 The percent of choices correctly predicted.....	106
Table 5.3 Dynamic probit models: alternative approaches	108
Table 5.4. Comparisons of the results reported in columns [1], [2], and [2] of	113
Table 5.3	113
Table 5.5 Initial conditions equations.....	114
Table 5.6 Dynamic probit models: average partial effects (APEs).....	116
Table 5.7 Estimated probabilities of participating (apes) for the presence of health problems, the presence of children, and state dependence for 2003, according to Heckman's (1981b) model.....	119
Table 5.8 Estimated probabilities of participating (apes) for the presence of health problems, the presence of children, and state dependence for 2003 according to Hyslop's (1999) model	119
Table 5.9 Hyslop's approach with correlated random effects (Mundlak's approach)	121

List of Figures

Figure 2.1 Individuals with very good or good self-perceived health, 2018 (% share of the persons aged 16 and over by sex).....	17
Figure 2.2 Individuals with long-standing (chronic) health problems, 2018 (% share of the persons aged 16 and over by sex).....	19
Figure 2.3 Women with long-standing (chronic) health problems, by employment status, 2018 (% share of the persons aged 16 and over)	20
Figure 2.4 Women in the uk with long-standing (chronic) health problems, by employment, 2018 (% share of the persons aged 16 and over).....	21
Figure 4.1 Employment and unemployment – non-employment rates from the ONS and from the sample.....	79
Figure 4.2 Self-assessed health by wave.....	93
Figure 4.3 Education by self-assessed health by wave	94
Figure 4.4 Self-assessed health by quintiles of non-labour income.....	95

“Different assessments of the unobservables have different effects on the interpretation of the evidence. For example, is joblessness due to unobserved tastes for leisure on the part of workers or a failure of the market to generate wage offers? Are women transients in the labour market or do some women (or most) have a long term attachment to it?”

Nobel Prize Lecture by James Heckman (2000, p. 265)

Chapter 1

Introduction

Over the last 50 years, the female employment rate has been continuously increasing in the United Kingdom (UK). According to the Labour Force Survey (LFS) in the UK, the employment rate among women aged between 16 and 65 years has risen to 72.5 percent in January 2020 from 52.8 percent in January 1971. In contrast, the employment rate for men has been declining over the same period. The male employment rate was 92.1 percent in January 1971 but fell to 80.4 percent in January 2020, although it remains consistently higher than the female employment rate. However, the gap between the male and female employment rates was just below 8 percentage points in January 2020, which is close to the lowest it has ever been since the Office for National Statistics (ONS) began recording these data in 1971. A decade earlier, the gap was 11.5 percentage points.

The key features behind the pronounced increase in female labour supply in the UK are changes in working patterns across the life cycle, rising educational attainment for women, diminishing marriage rate or cohabitation and increasing divorce rate, and the decrease in childbirth.¹ All these changes reflect greater social acceptance of women in the labour market through emerging and diverse working arrangements, including increase in part-time employment and increase in workplace flexibility, such as working from home.

According to the utility maximization framework of consumer theory and time allocation, “market wage” offered and the “reservation wage” (or the “shadow price of time”) are the key determinants of whether a woman decides to participate in the labour force, under a budget constraint. Almost all recent studies, however, disembark from this narrow framework and recognize that women’s criteria upon participation vary over the life cycle and differ from woman to woman. For most women, the decision to participate or not is influenced primarily by financial incentives, particularly when they possess suitable skills and training that are appealing to employers. For other women, the decision

¹ Roantree and Vira (2018) provide an excellent summary of the emerging working patterns for working-age women in the UK from 1975 to 2017 and compare cohorts for women born in 1945, 1958, 1960, 1970 and 1980, using data from the Labour Force Survey and the Family Expenditure Survey.

to participate is driven principally by personal circumstances, personal characteristics, personal preferences or taste shifters, and the self-esteem that employment brings. Nevertheless, in all cases, women weigh the benefits and costs of allocating time between working, leisure, domestic duties and childcare for their decision to participate.

The patterns of female labour supply over the life-cycle have been exhaustively documented during the last fifty years in economics and sociology. For example, Hakim (1979) examined the changes in women's labour force participation over their working lives. She calls the pattern of women's participation a "two-phase" working life. It is a two-phase phenomenon because women have high rates of employment when they first leave school or higher education, their rate of employment then declines, but it begins to increase again after the 35 years of age, reaching a peak in their 40s. Dex (1987) has named these two phases of employment "the family formation phase" and the "final formation phase." During both phases, women may have a range of employment profiles including returning to work between births.

The literature explored extensively possible determinants of female labour supply, either with the use of cross-sectional data or panel data. Most studies have shown that the key determinants mostly are marital status, presence of children, wages, non-labour income, spouse's employment state, years of education, age, and years of work experience. Those most important factors that reduce the probability of employment are the presence of young children and high levels of alternative income. On the other hand, education and work experience are the most important factors that increase employment.

More recently, Eckstein and Lifshitz (2011) distinguish the dominant trends in female employment over the last 50 years in five categories; the increase in women's education, the increase in women's own earnings and the consecutive narrowing of the gender wage gap, the decrease in women's fertility, the decrease in the marriage rate, and sequentially the increase in the divorce rate and, finally, "other" factors that are hard to measure precisely, such as factors that include technological evolution that affect time that women spend at home and changes in social ethics and morals.

Early in the 1980s, a particular interest was in the labour supply of married women because they constituted the vast majority of working women since the end of World War II and, also, because they do not work during significant periods of their life cycle.

In addition to the above determinants, extensive research on married women's labour supply with the use of panel data has repeatedly indicated that labour supply decisions of married women are strongly characterized by intertemporal persistence over time; see, for example, Heckman and Willis (1977), Nakamura and Nakamura (1985), and Eckstein and Wolpin (1989). The term 'intertemporal persistence' means that women are inclined to stay in the same employment state for most of their working life, whether this is employment or non-employment.

Three are the key elements of intertemporal persistence in the female labour supply over the life-cycle: state dependence, individual heterogeneity, and serial correlation. The first (state dependence) indicates whether and to what extent the previous participation state affects the probability of current participation, after taking into account the initial conditions and controlling for observed and unobserved individual heterogeneity. State dependence (also referred to as "true" state dependence by Heckman, 1981a) is generated by two mechanisms. The first is that employment may be associated with accumulation of human capital and non-employment may be associated with a decline and potential depreciation of human capital. The presence of state dependence is an important determinant in research that examines the effect of fertility/presence of children on female labour supply. For example, if a woman interrupts her work to give birth and raise her newborn, this may cause a decline in her human capital stock because she misses work experience and training opportunities. Consequently, the decline of human capital stock may make her return to employment quite difficult, especially if the labour market does not provide flexible work arrangements.

The other mechanism that may generate state dependence is that employment and, more frequently, non-employment may be associated with increased job search cost. For example, an unemployed individual may have more difficulty in finding work compared to an employed individual. Such job search costs may make women, who were employed before they gave birth, to decide to stay in their employment after a maternity leave, thus inducing persistence of work. Therefore, we may observe a high persistence in employment even if the effect of fertility is strong. The two mechanisms could also be considered as indicators of labour market rigidities. For example, apart from capital accumulation and job search costs, Edon and Kamionka (2010), Del Boca and Sauer (2009) and Michaud and Tatsiramos (2008) give evidence that the differences in the degree

of state dependence in a cross country framework can also be due to childcare institutional factors, such as availability of childcare provision and child benefit policies.

The second key element of intertemporal persistence in the female labour supply is unobserved individual heterogeneity (also referred to as “spurious” state dependence by Heckman, 1981a). Heckman described unobserved individual heterogeneity by stating that “individuals may differ in certain unmeasured variables that influence their probability of experiencing the event, but that are not influenced by the experience of the event.” Hence, persistent individual heterogeneity encompasses observed and unobserved individual characteristics, which are constant over lengthy phases of the life-cycle and which make a woman more or less likely to be employed, irrespectively of her work history. Observed characteristics might be the educational qualifications. Unobserved characteristics might be attributes such as motivation, inner abilities, and intelligence (or lack of them). For example, a woman with high propensity in maternity has weaker preference for employment, whereas a career-oriented woman has stronger preference for employment. This difference in preferences for employment generates persistence because of self-selection of women with strong preference for employment and those with weak preference for employment. This difference in preferences arises from unobserved heterogeneity, which may lead to the possibility of the so-called spurious state dependence.

The distinction between state dependence and unobserved individual heterogeneity is very important because unobserved individual heterogeneity may make an individual more likely to be employed and this mistakenly may lead us to the conclusion that true state dependence is present when in fact it is the time-invariant individual characteristics that increase the likelihood for a woman to be employed in the current year if she was employed in the previous year. It is therefore important to disentangle the two potential elements of intertemporal persistence because they have different policy implications. For example, social and economic policies targeted to prevent unemployment would have a much more significant effect if there is positive state dependence in unemployment than targeted to retrain unemployed individuals. On the other hand, if there is evidence that unemployment has a propensity to replicate itself due to individuals’ characteristics, policies targeted to create an incentive for individuals to return to the labour market may be more appropriate.

State dependence is measured econometrically by introducing the response variable lagged one period as a regressor. According to Heckman (1978) (cited in Nakamura and Nakamura, 1985), a dummy variable that equals one if a person worked in the previous year and zero otherwise could serve as a good proxy for unobservables that affect the work behavior of a person year after year (page 281). Similarly, Nakamura and Nakamura (1985) have strengthened this argument with the following statement: “Incorporation of information about past work behavior, even in the form of a simple dummy variable is found to result in greatly improved forecasts of the employment and earnings behavior of wives over time” (page 291) as well as by the statement “...the reformulation of structural models so that lagged dependent variables appear as explanatory variables in these models may prove to be one of the more tractable ways of controlling for fixed and persistent unobservables.” A positive coefficient of lagged participation status (or lagged hours of work) indicates that past participation status (or past hours of work) has a positive impact on current and future participation status (or current and future hours of work). The larger the coefficient of the lagged response variable, the more persistent the labour supply will be over time.

Unobserved individual heterogeneity is measured econometrically by introducing an individual time invariant effect into the regression. A high individual effect implies that the woman tends to stay in labour force participation, whereas a low effect indicates a propensity to stay out of participation. In other words, large individual effects imply a more persistent female labour supply. In addition, omission of individual heterogeneity from a labour supply model would lead to bias upwards the coefficient of state dependence, because unobserved individual heterogeneity is potentially correlated with the lagged response variable (Shaw, 1994). Thus, unobserved individual heterogeneity is a key determinant of female labour supply.

Finally, the third key element of intertemporal persistence in female labour supply is serial correlation, which is measured econometrically by allowing unobserved time-varying transitory shocks in the regression to follow an AR(1) process. Both Hyslop (1999) and Keane and Sauer (2009) highlight the importance of serial correlation as an important element of persistence. For example, an individual who has experienced health deterioration in the previous year may be more likely to experience health deterioration in the current year as well. Keane and Sauer (2009) claim that serial correlation serves as an

indicator of persistence in shocks to tastes and/or productivity, and Hyslop (1999) as a more general form of state dependence. For example, if a health indicator is included in the regression as an explanatory variable, a positive correlation of health deterioration over time may be reflected in positive correlation between deterioration of health and non-participation in the labour supply model (Cai and Kalb, 2006; Laplagne, Glover and Shomos, 2007).

Besides the above determinants of female labour supply, an emerging but well-established finding in the last thirty years is that health state has a significant impact on the labour supply of all individuals, men and women. Chirikos (1993) in his comprehensive overview, states:

In virtually every case, impaired health exacts some toll by either restricting the ability of individuals to engage in market work or shifting their preferences for time spent in the labour market, reducing the wages of workers in poor health, and/or changing the labour market behavior of other persons in the household of the health-impaired individual. This conclusion is generally invariant to sociodemographic characteristics, occupational or industrial attachment, or the type of physical or mental health condition (page 301).

This passage suggests that not only does deterioration of health reduce the labour supply of individuals, but moreover, it is an important factor that leads individuals out of the labour market and reduces their probability of entry or re-entry into employment. Less healthy individuals may not be able to work as effectively as those with better health and may find it more difficult to find suitable work. They are also more likely to be less productive and accept a lower wage, which may lead them to decide to drop out of the labour force. Finally, less healthy individuals may have more difficulty in finding a job and thus may feel discouraged to continue searching, even though they want to work. In an extension of this point, Oguzoglu (2010) argues that “persistence in employment can mask or overemphasize the real impact of a work limitation. For example, a disabled person's failure in the search for a job may be due to previous search failures rather than to the disabling conditions themselves.” Thus, the incorporation of health into labour supply models becomes necessary if we want to avoid the omitted variable bias in our estimation. In addition, the interpretation of the effect of health-related variables should take into account the persistence of the participation status.

The springboard for researchers to examine the interaction between health and labour market outcomes (wages, hours of work, labour market attachment, and retirement) has been the concern in the economic impact of the ageing of the population due to early retirement decisions of workers because of health-related problems. Age has been considered as the key determinant of the interaction between health and labour supply in these researches, because as age increases health deteriorates and labour supply diminishes (see, for example, Haan and Myck, 2009).

According to the human capital theory, health is treated as a form of human capital (Becker, 1965). Indeed, individual health capital has many similarities with education. Health, like education, to the extent that individuals are physically and mentally able to perform their work, increases the productivity of workers and raises the wage. Hence, health has a significant impact on labour supply, and health and labour supply are positively correlated. This means that healthier individuals are more likely to be in the labour market.

The human capital approach by Becker was extended by Grossman (1972) who established a pioneering economic model of demand for health by individuals. In his model, Grossman (1972) considers health as both a consumption and investment commodity and, unlike other forms of human capital, which need not be depreciated, the initial stock of health must be constantly replenished by sacrificing time and monetary resources. This means that individuals need to invest in their health if they want to maintain or improve their current health state. This is presumably achieved by investing in health by means of time and money (Breunig, 2011, Chapter 1, p. 3). For example, individuals might need to spend more time for relaxation, recreation or exercise and spend more income to cope with increasing medical expenses. In turn, the availability of time and monetary resources may depend on the individual's labour supply, past and current (Cai, 2021). Therefore, "while earnings are partly determined by investments in health capital, the stock of health today depends on past (and current) investments in health, which, in turn, depend upon past (and current) earnings" (Breunig, 2011). This implies that individuals decide upon their labour supply simultaneously with health, as they would do with other commodities. This, in turn, implies that theoretically health may be endogenous to labour supply.

Grossman (1972) gives the benchmark to consider the possibility that the relationship between labour supply decisions and health status may be more complex. Although human capital theory argues that deterioration of health decreases the productivity of workers, their wage and their attachment to work, it may also be suggested that decreased wages associated with deterioration of health might increase hours of work which could be demonstrated by including an interaction term of health and wage with a negative coefficient. Cai and Kalb (2006) argue that individuals may increase their labour supply as their health deteriorates, because they need more income to encounter their medical expenses. This means that less healthy individuals presumably need to increase their labour supply in order to finance the expenses of their increased demand for health care services. Further, Cai (2021) argues that the onset of a health problem, for instance, may lead people to value time out of the labour market more as the time needed to care for one's health increases with ill health. Additionally, if deterioration of health impacts life expectancy, it may lead individuals to withdrawal from the labour market or, if eligible, to early retirement. Thus, theoretically, the precise direction of the health effect on labour supply is rather ambiguous, but most empirical studies find that a deterioration in health reduces labour supply.

Beyond the considerations based on Grossman (1972) for the endogenous treatment of health in labour supply, another source of the endogeneity of health is that there may be a feedback of work on health, whereby employment affects directly health (Currie and Madrian, 1999). For example, working in a stressful environment or being unemployed for a long period may have detrimental consequences for mental health (Clark and Oswald, 1994). This, in turn, may further worsen an individual's prospect to entry the labour market. Haan and Myck (2009) explore the relationship between non-employment and health by applying panel data estimation in which they allow for unobserved individual heterogeneity. Their results demonstrate a significant interaction between non-employment and deterioration of health and persistence in both processes. Delattre, Moussa and Sabatier (2019) investigate the association between health and employment status within a Granger causality framework in which they take account for state dependence, initial conditions, and unobserved heterogeneity. Their results emphasize the importance of all these three elements and reveal Granger causality (i.e., predictability) between health and employment status. Nonetheless, this is not always a consistent result in empirical studies. For example, while Lindeboom and Kerkhofs (2009) find a strong

impact of health on employment status, they do not find the same for the impact of employment status on health of men in the Netherlands. On the other hand, good social interactions at work, personal fulfillment from employment engagement and the self-esteem that employment brings may benefit health (Cai and Kalb, 2006). Thus, employment status could also affect health, although the direction of the impact is also ambiguous. Nonetheless, regardless of the empirical validity for the dual causality between employment status and health, the possibility further suggests that health should be treated as endogenous (Currie and Madrian, 1999).

Consequently, the interaction between health and labour supply can work in different directions. In the majority of studies, deterioration of health reduces labour supply, whereas labour supply can also impact on health, although the direction of the impact is ambiguous. Nonetheless, regardless of the complexity of the interaction between health and labour supply, it becomes apparent that health capital cannot be treated as exogenous in labour supply models. Rather, it must be treated as a choice variable (Currie and Madrian, 1999, p.3312).

On the other hand, Currie and Madrian (1999) also review the issue of exogeneity of health-related variables. They find that most of the literature surveyed until 1999 treats health as an exogenous variable. This is strongly justified by the fact that exogenous health shocks are the main factor that can cause variation in health status, at least in developed countries. Currie and Madrian (1999) argue that this may not be an unreasonable assumption, given that current health depends on past labour supply decisions and habits with great persistence that may be very difficult to break, such as smoking or high fat diets. Finally, Currie and Madrian note that the assumption of exogeneity of health is often based on the fact that individuals often have imperfect knowledge about their health “production function” at the time the labour supply decision is made (p. 3113).

Another important issue in the literature is that health is a multi-dimensional concept and its definition depends in part on the questions the researchers address. Furthermore, there is no consensus about the measurement of health that is more accurate as an indicator for health. Consequently, both the definition of health and the measurements used in the empirical literature vary from study to study. A great part of studies uses subjective measurements, such as self-assessed health, health limitations to daily activities or health limitations to work as proxies for health. Other studies focus on

more objective physician-diagnosed measurements of health. Whatever the measurement is, the concern is almost always about the degree to which these measurements fully encompass the multiple dimensions of health that may affect productivity and labour supply. As Currie and Madrian (1999) highlight in their comprehensive overview, "...health has a pervasive effect on most outcomes of interest to labour economists including wages, earnings, labour force participation, hours worked, retirement, job turnover, and benefit packages. But, unfortunately, there is no consensus about the magnitude of the effects or about their size relative to the effects of other variables" (pages 3310-3311).

As we mentioned earlier, the existing literature of female labour-force participation concentrates on marital status, presence of children, wages, non-labour income, spouse's employment state, years of education, age and years of work experience as it has been argued that these factors have the greatest impact on women's work decisions, both theoretically and empirically. This thesis contributes to the existing literature by focusing on how women's work decisions vary with the insertion of health-related measurements for physical and mental health among the other factors. There are a few studies that investigate specifically how health affects women's labour supply. For example, Arber, Gilbert and Dale (1985) used cross-section data to explore the interaction between employment status and health. Majeed, Forder, Mishra and Byles (2015) used panel data to examine the impact of chronic diseases (diabetes, asthma, depression and arthritis) on employment status for middle-aged women, and a few other researchers have included only a binary indicator of health in a cross-country comparison (Michaud and Tatsiramos, 2005 and 2008). Our framework is determined by Heckman (1981b), Hyslop (1999), Keane and Sauer (2009) and Bartolucci and Nigro (2010) for dynamic discrete choice models. Heckman and Willis 's (1977) study prompted these models, because it became necessary to distinguish between "true" state dependence, captured by the effect of the coefficient of the lagged dependent variable, and "spurious" state dependence, caused by the presence of individual time invariant characteristics, when analyzing the transitions in and out of work. Another issue that arises in dynamic specifications of non-linear panel data models (where a dependent variable is assumed to depend on its past value) is the initial conditions problem: the observed start of the examined period does not necessary coincide with the true start of the stochastic participation process. As a result, the participation state at the first year of observations cannot be considered exogenous to

participation (the dependent variable) because there might be correlation between the lagged dependent variable and the individual heterogeneity. Heckman (1981b) tackled the initial conditions problem by modelling the joint distribution of the outcomes conditional on participation state in the first year of the sample. That is, Heckman (1981b) proposed a solution which involves a simultaneous estimation of two equations: a “structural equation” and a “reduced-form equation.” For the years following the initial year, the structural equation is used to estimate the probability of participating in the current year as a function of the participation status in the previous year. The reduced-form equation accounts for the initial conditions and is used to estimate the probability of participating in the first year of the sample. Heckman (1981b) solves the initial conditions problem, accounts for unobserved individual heterogeneity, and distinguishes “true” state dependence from “spurious” state dependence. Hyslop (1999) extended Heckman (1981b) and attempted to isolate the true state dependence from both individual heterogeneity and autocorrelation in the error term. That is, the persistence in the time-varying error term is also parameterized. Keane and Sauer (2009) further generalized Hyslop (1999) and introduced a more flexible treatment of the initial conditions equation (i.e., the “reduced-form equation”) in which they defined an additional parameter. Bartolucci and Nigro (2010) tackled the initial conditions problem via a fixed-effects approach and estimated the coefficients of the regression consistently without imposing distributional assumptions on the unobserved heterogeneity, which is eliminated by using a suitable sufficient statistic. Nevertheless, a considerable pitfall of this approach for dynamic panel data analysis is that the model identification is based on those observations where the dependent variable (participation status) changes over time for an observation to contribute to the maximum likelihood estimation. The result of this pitfall is the drastic reduction of usable observations. We, therefore, cannot explore effectively the assumption that the unobserved individual heterogeneity is correlated with the explanatory variables and lagged participation. Nevertheless, we use Bartolucci and Nigro’s estimator (2010) in an attempt to investigate female labour supply in a conditional fixed-effects approach. Finally, the specification of Mundlak (1978), embedded in Hyslop’s (1999) approach, allows the unobserved individual heterogeneity to be correlated with the explanatory time-varying variables and specifies the relationship between them through a linear form.

Our sample is constructed with the use of the first 12 waves (for the years 1991-2002) of the British Household Panel Survey (BHPS) in the UK and consists of 331 women

aged between 39 and 45 years old, following Eckstein and Wolpin (1989) for the choice of this age range.

The existing literature of the impact of health on labour supply, gives its emphasis either on decisions to retire or on work decisions of men, until nowadays. Currie and Madrian (1999, page 3353) emphasize that “a glaring limitation of the existing literature is the intense focus on elderly white men, to the virtual exclusion of most other groups. Studies to remedy this situation would be most useful.” The choice of the age group and the span of twelve years eliminates, hopefully, the possibility of retirement, because we want to explore women’s work decisions, excluding decisions to retire. Therefore, this thesis attempts to fill this gap in the literature by focusing on women of working age.

To the best of our knowledge, our two contributions to the literature are as follows: (1) it is the first study that examines the impact of both physical and mental health on women’s work decisions as we use several aspects of health as proxies for health (in particular, self-assessed health, whether health limits daily activities and health problems related to anxiety and depression), and (2) it is the first study that examines female labour supply within a framework determined by Bartolucci and Nigro (2010).

From a policy perspective, increasing female labour participation is a significant factor that contributes to the economic development, especially in an ageing society. Identifying the determinants of work decisions of women and, in principle, the barriers that women are likely to face over the life course, which are hidden behind the labour supply decisions, ameliorates and strengthens our understanding of the dynamics of their labour supply and its interaction with the economic environment.

At the same time, an ageing society imposes economic challenges for the economy. A principal challenge is a continuously increasing demand for public expenditure to cope with the pensions for the retired individuals and another one is the increasing demand for human resources to support national production, part of which is intended for paying the pensions and similar benefits. Therefore, in order to address the economic challenges of an ageing society, the increase in employment rates are considered fundamental by the policy makers. In this context, policy makers also highlight the importance of health in the labour market because they recognize its effect on productivity and labour supply in general, but also acknowledge the cost of health in terms of productivity loss due to early

retirement decisions. For middle-aged women, there already exist policies aimed to engage women in paid employment and encourage those who do not work to enter the labour market, despite their personal circumstances and family duties. Nevertheless, for women who have health problems and for those who have experienced a health shock such policies are limited. There could be policies that promote uninterrupted employment trajectories for women despite health problems, including greater workplace flexibility to respond to their need for medical care, stress-free work environments for ill women and also, equally important, public actions targeted to prevent disease. Besides, a greater childcare provision for these women should be of a prime concern.

In sum, the effectiveness of policies designed to support simultaneously labour force participation and health of women depends on a better understanding of the impact of health on labour supply and on the complex nature of the relationship between labour force participation and health. Such policies could also be applied for other disadvantaged and vulnerable population groups.

The rest of the dissertation is organized as follows. Chapter 2 consists of three parts. The first part presents trends and patterns in female labour supply and health in the EU and in the UK with data from Eurostat. The second part reviews the literature on female intertemporal labour supply with a focus on studies that employ dynamic discrete choice models in order to explore state dependence, unobserved individual heterogeneity and serial correlation. The third part reviews the literature on the link between labour supply and health, concentrating again on studies within an intertemporal context. Chapter 3 presents the theoretical framework within which women decide to participate in the labour market or not, as well as the econometric approaches. Chapter 4 describes the construction of our sample and the summary statistics of our dataset. Chapter 5 presents and discusses the estimation results and chapter 6 concludes.

Chapter 2

Review of literature

2.1 Introduction

Almost all recent studies disembark from the utility maximization framework of consumer theory and time allocation where the “market wage” offered and the “reservation wage” (or the “shadow price of time”) determine whether a woman decides to participate in the labour market or not. Rather, they adopt the well-established conception/view that personal characteristics, personal preferences or taste shifters and personal circumstances over the life-cycle also determine labour supply decisions of women. These findings, together with the fact that individuals are aware of personal preferences but the econometrician only observes market wages and hours worked, have led almost all labour supply studies to employ the “reduced-form” approximation within a human capital framework in which market wages and reservation wages² are substituted by observed personal characteristics. In contrast, the structural approximation³ considers that the individuals choose their optimal combination of consumption and leisure over the lifetime in reference to the “market wage” offered and the “reservation wage” and its aim is to estimate the parameters of the utility function and the resulting labour supply function. In practice, structural models are derived with the use of dynamic programming methods which are extremely demanding in the computational apparatus.

Section 2.2 presents trends and patterns in female labour supply and health in the EU and in the UK with data from the Eurostat. We include this section with graphs in our review because we believe it initiates us in the subsequent sections. In section 2.3, we

² Accordingly to Heckman (1974), reservation wage or “shadow price of time” of married women is the value of time for non-working women and indicates the wage upon which a woman decides to participate in the labour market or not.

³ See, for example, Keane, Todd and Wolpin (2011).

review the literature on female intertemporal labour supply, and particularly on women's labour supply decisions; whether a woman chooses to participate in the labour market or not (and the hours she works, in a few researches) across time. We do not attempt to provide a review of the extensive female labour supply literature.⁴ Instead, this thesis is built around empirical studies and the main interest is in the intertemporal persistence caused by state dependence, individual heterogeneity and serial correlation in the error terms. Despite the extensive literature on female labour supply, only a few studies have investigated the intertemporal female labour supply for which dynamic discrete choice labour supply models are estimated with the use of panel data. By “dynamic” we mean that we investigate the transitions⁵ into and out of labour force participation of women. In empirical analysis, this is achieved by incorporating the lagged variable of participation decision (whether a woman chooses to participate or not is identified by the employment status in the data) among the explanatory regressors. Our attention is on “reduced-form”⁶ approximations, but we also give a brief literature review on structural approximations in subsection 2.3.2.

We begin section 2.4 with definitions of health, measurements of health and the potential endogeneity of health in labour supply models. Subsections 2.4.3 and 2.4.4 review the literature on the link between labour supply and health, specifically on the impact of health on participation status. With a few exceptions, we give our emphasis on studies that use panel data and review dynamic models because they give us the opportunity to explore intertemporal persistence in labour supply. Further, where available, we comment on results related to women as our purpose is to link this review with the review of section 2.3.

⁴ For a comprehensive review, see, for example, Killingsworth and Heckman (1986), Blundell, MaCardy and Meghir (2007) and Myck and Reed (2005).

⁵ See, for example, Myck and Reed (2005) for a comprehensive review of models of labour market transitions (2005).

⁶ See, for example, Heckman and Willis (1977).

2.2 Figures and trends

Data from the Eurostat reveal that more than two out of three people in the EU perceived their health as very good or good and more than one out of three aged 16 years or over reported having a long-standing illness or health problem in 2018.

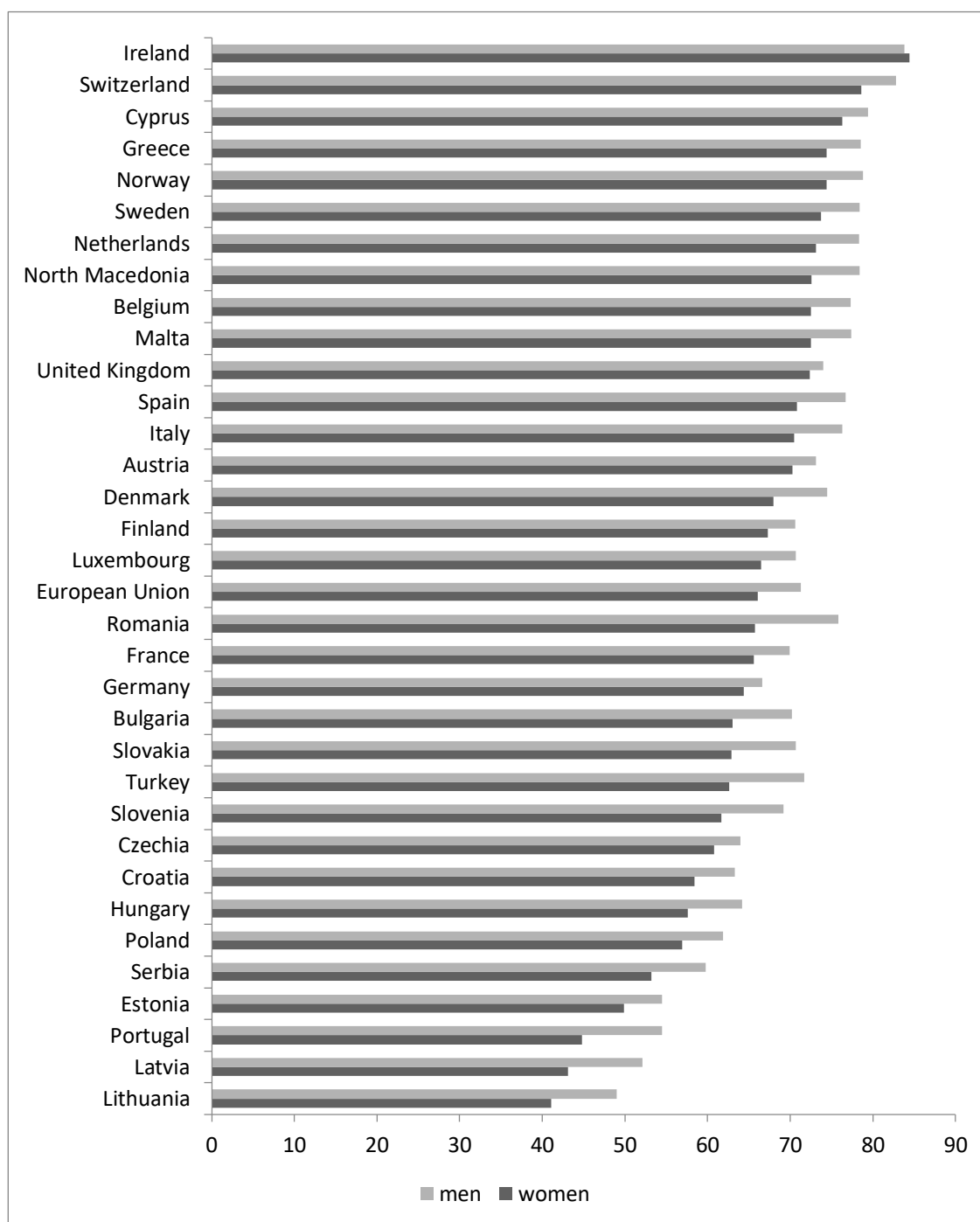
For the measure of self-assessed health, in particular, Figure 1 shows that in 2018, about 68% of the European population aged 16 and above reported having very good or good health. The highest proportions were reported in Ireland (84.1%) and Switzerland (80.7%) and the lowest ones were reported in Latvia (47%) and Lithuania (44%). For the UK, the percentage for this measure was 73.2%. From Figure 1, there also appears to be a gender health gap.

Figure 2.1 also demonstrates that men tend to rate their health better than women. In 2018, men were more likely to rate their health as very good or good than women in all EU Member States, except for Ireland, where the percentage of women reporting very good or good health exceeded that of men (the gender health gap was 0.6 percentage points). By this measure, the largest gender health gaps were recorded in Romania (10.1%) and Portugal (9.7%) and the smallest in Germany (2.2%) and the UK (1.6%). Across the EU-27 as a whole, the gender health gap was slightly more than 5 percentage points, as 71.3% of men rated their health as very good or good compared with 66.1 of women.

Similar to self-perceived health, Figure 2.2 reveals that more than one out of three people in the EU aged 16 years or over reported having long-standing health problems in 2018 (36% of the total population in the EU) with the percentages being 37.9 and 34.1 for women and men respectively. From Figure 2 we notice that in all countries, with the exception of Ireland only, the percentage of women who declare a long-standing health problem exceeded the corresponding percentage of men. The largest health gap was observed in Turkey (10.5%), Lithuania (10%) and Latvia (9.4%) and the narrowest health gap was in France and Malta (both 2%). Among the EU Member States, the largest proportions of women reporting having long-standing health problems were observed in Finland (52.4%) and Estonia (50.7) and, by contrast, the smallest proportion was in Italy (17.2%). In the UK, 43.2% of women reported long-standing health problems compared to 40.2% of men.

Looking at women aged 16 and above, there reveals a relationship between employment status and the presence of long-standing (chronic) health problems (see Figure 2.3). Whereas 28.3% of employed women in this age range in the EU reported such problems in 2018, the proportion was 47.8% for non-employed women.

Figure 2.1 *Individuals with very good or good self-perceived health, 2018 (% share of the persons aged 16 and over by sex)*

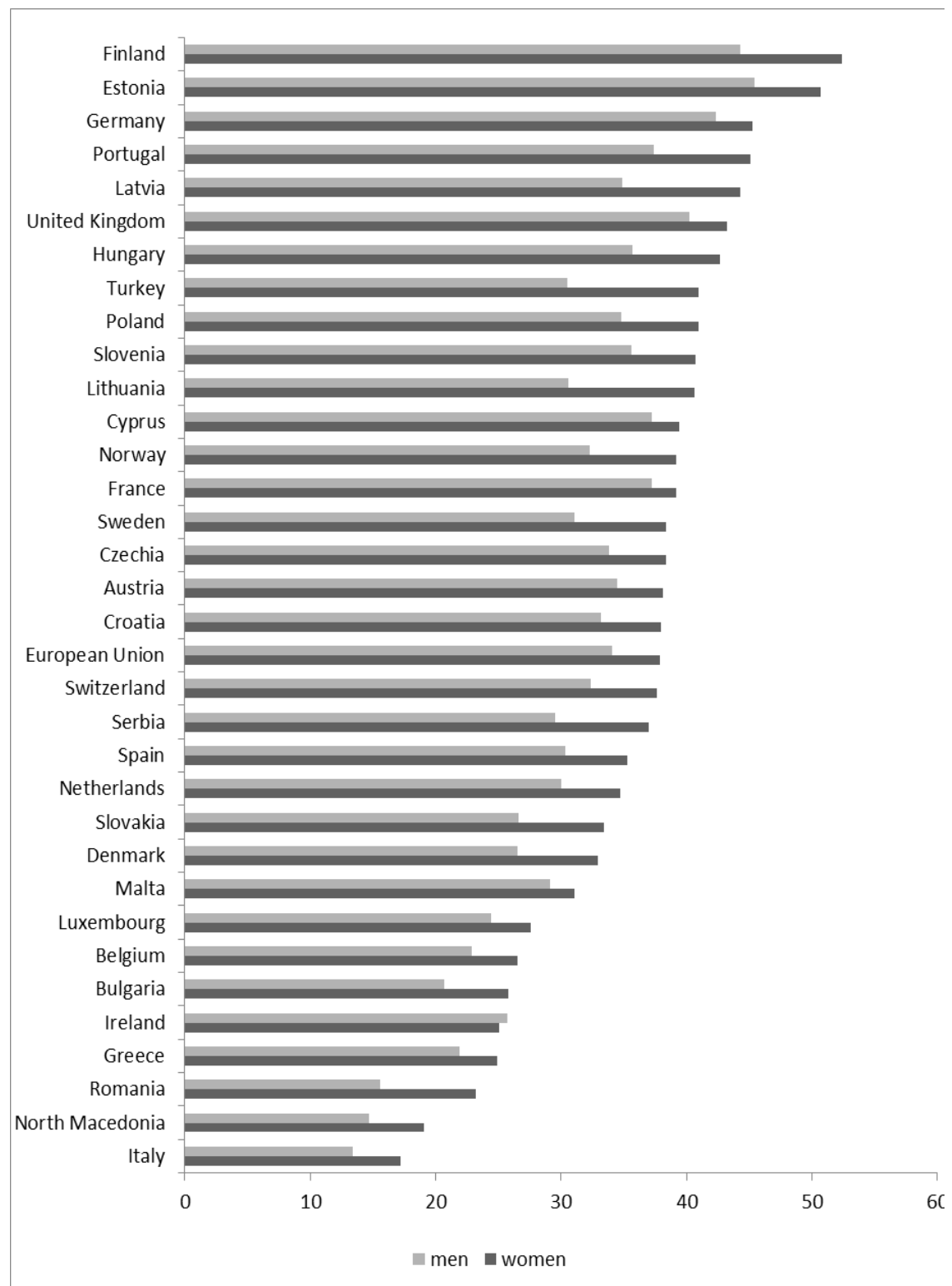


Source: Eurostat ([hlth silc 10](#))

All countries reported the same broad pattern for the working-age population of women, with a smaller percentage of women reporting long-standing health problems among employed women than among non-employed women. In percentages, the largest difference was reported for Lithuania, the proportion of 61% for the non-employed women and 20.4% for employed women resulting in a difference of 40.6 percentage points (p.p.) Other countries where the difference exceeded 30 percentage points included Latvia (33.50 p.p.), Serbia (32.70 p.p.) and Croatia (32.30 p.p.), followed by Austria (24.4 p.p.), Sweden (23.8 p.p.) and Slovenia (21.2 p.p.). The lowest difference, 20 percentage points or less) in the presence of long-standing health problems between employed and non-employed women were reported in Turkey (14.20 p.p.) followed by Denmark (15.60 p.p.), Switzerland (16.50p.p.), Italy (17.20 p.p.) and Luxembourg (17 p.p.). For the UK, the proportion of women having long-standing health problems is 32.6% and 57.1% for the employed and non-employed women, respectively.

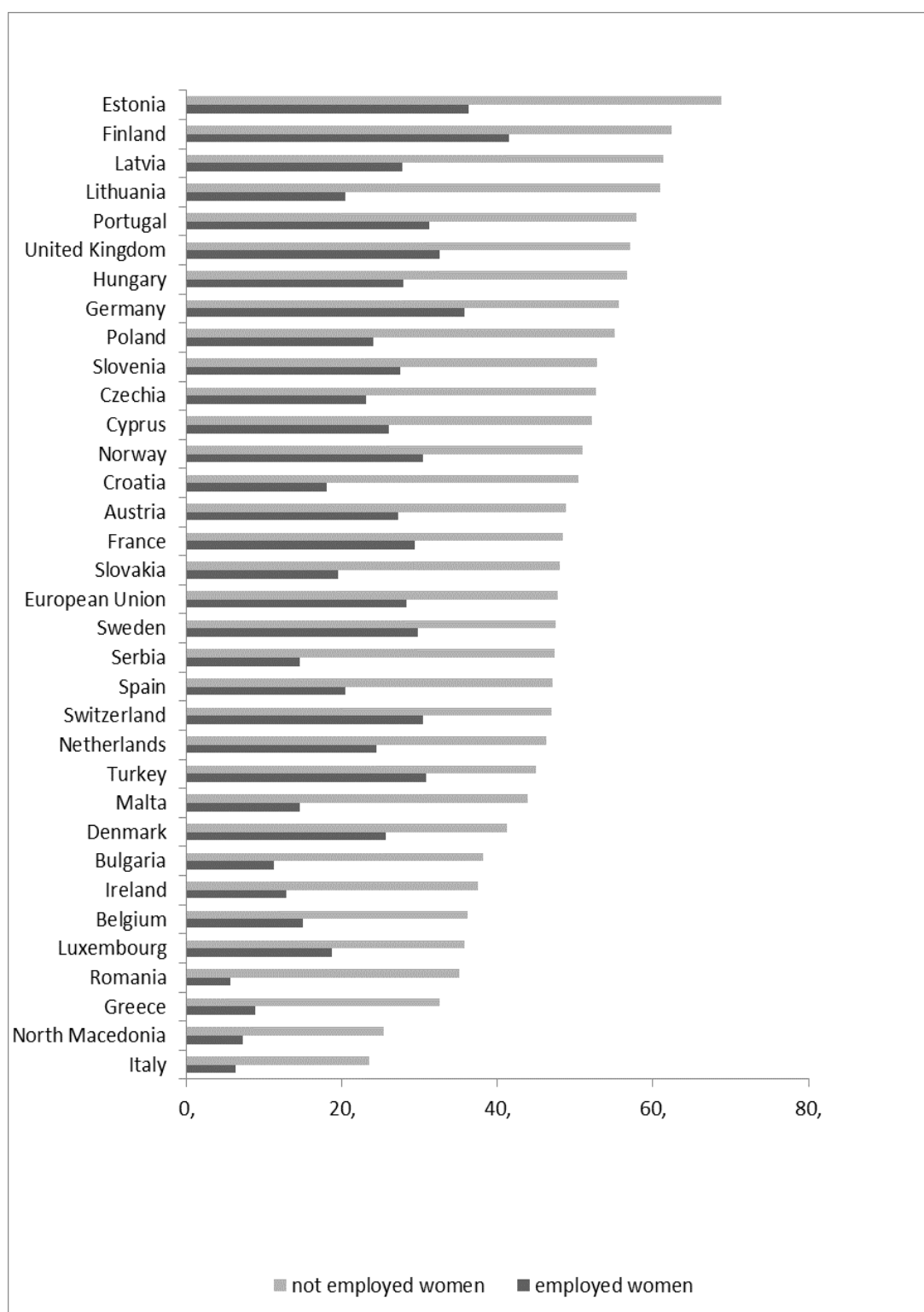
An interesting feature is that all countries exhibit two-digit numbers for the percentage of women having long-standing health problems and being employed, except for Romania (5.6%), Italy (6.3%), North Macedonia (7.2%) and Greece (8.9%). An explanation for this, among other factors, might be the absence of work-related policies that could encourage women to participate or remain in the labour market.

Figure 2.2 Individuals with long-standing (chronic) health problems, 2018 (% share of the persons aged 16 and over by sex)



Source: Eurostat ([hlth_silc_11](#))

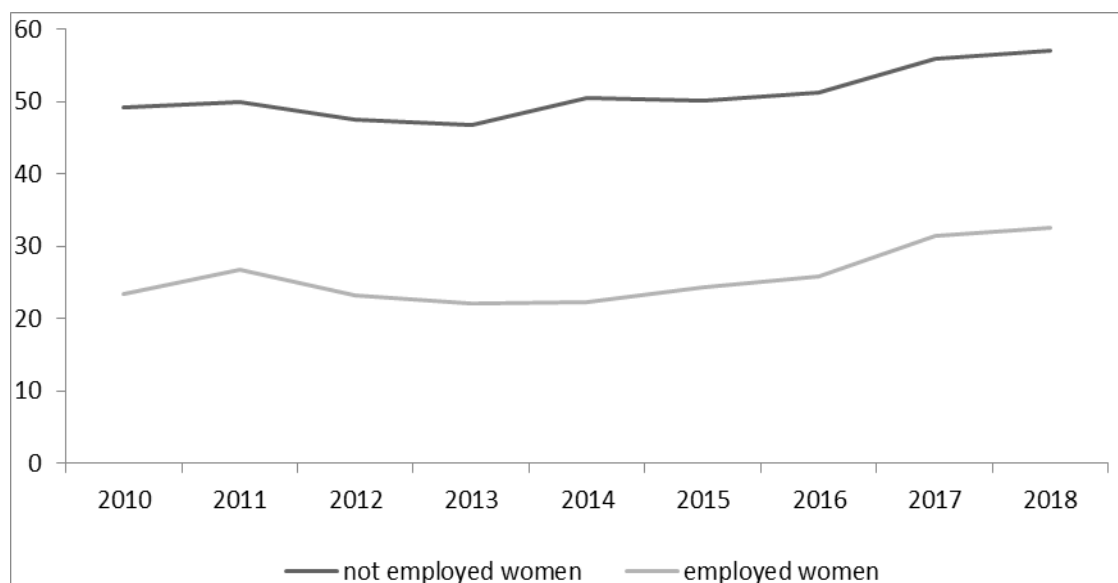
Figure 2.3 Women with long-standing (chronic) health problems, by employment status, 2018 (% share of the persons aged 16 and over)



Source: Eurostat ([hlth_silc_04](#))

In figure 2.4, we can see the trend of employment status for women who report having long-standing health problems from 2010 to 2018, in the UK. The figure reveals an increasing proportion of both employed and non-employed women across time, especially in 2013 and afterwards. The largest difference between the two proportions was observed in 2011 (23.1 p.p) and the highest percentage of non-employed women who report having long-standing health problems was in 2018 (57.1%).

Figure 2.4 Women in the UK with long-standing (chronic) health problems, by employment, 2018 (% share of the persons aged 16 and over)



Source: Eurostat ([hlth_silc_04](#))

2.3 Women's work decisions

2.3.1 Reduced-form approximations in female labour supply studies

An early research on female labour supply was conducted by Heckman and Willis in 1977. They concluded that women may have become permanent workers because work experience showed women the need for working in order to develop and retain their investment in human capital. This further led them to the conclusion that "...of otherwise identical women, those women who worked in year $t-1$ are more likely to work in year t than are those women who did not work in year $t-1$ " (Heckman and Willis, 1977, p. 17).

Four years later, Heckman (1981a) explored married women's labour supply decisions using the first three waves from the Michigan Panel Survey of Income Dynamics (PSID). His sample consisted of women aged 45-59 years old and employed various probit models to investigate whether or not recent work experience increases the probability that a woman will participate in the labour market in the future and to explore whether controlling for unobserved individual heterogeneity in panel data impacts this probability. Heckman found that a dynamic model that accounts for individual heterogeneity performed better than other model specifications. Other specifications that ignore individual heterogeneity overstated the effect of previous recent work experience. Heckman (1981a) stressed the importance of recent work experience in his findings stating that "it is often noted that individuals who have experienced an event in the past are more likely to experience the event in the future than are individuals who have not experienced the event. The conditional probability that an individual will experience the event in the future is a function of past experience" (p. 91). This is state dependence, or "true" state dependence, for emphasis. Heckman (1981a) also described state dependence in this article as "past experience has a genuine behavioral effect in the sense that an otherwise identical individual who did not experience the event would behave differently in the future than an individual who experienced the event" (p. 91). In other words, state dependence means that the experience of participating contributes to choices relevant to future labour supply decisions to be altered or, equally, that the participation state in which a woman currently is, changes the probability of the participation state she will be in the future.

Nakamura and Nakamura (1985) strengthened the importance of using adequate panel data into the investigation of the determinants of female labour supply. Their incentive was the awareness of the need for a deeper understanding of the origins of the rising labour supply of married women in the United States and Canada since World War II.⁷ Using panel data⁸ from the Michigan Panel Study of Income Dynamics (PSID) from 1969 to 1978, they extended Heckman's approach and, in addition to the binary lagged dependent variable, which is whether or not a married woman worked in the previous year, they incorporated hours of work and female earnings in their model. Their findings confirmed Heckman's findings with respect to years of work and non-work. That is, the

⁷ See McCarthy Helen (2016) for married women's employment in Post-War in the UK.

⁸ The sample consisted of 546 continuously married women aged between 21 and 64 years old for whom information was available for 10 consecutive years, from 1969 to 1978.

percentage of women who work in the current year is much larger for those who worked in the previous year compared to those who did not. In addition, among women who have worked in the previous year and work in the current year, a greater proportion (of those who work in the current year) are those who worked more hours and earned higher hourly wages in the previous year, compared with those who worked fewer hours and had lower earnings in the previous year. They reached the following conclusion: “We believe it would be important if researchers could identify what observable factors, if any, increase the likelihood that wives will alter their work behavior from what it has been in the immediate past, even if we are not able to fully understand or explain this previous work behavior” (Nakamura and Nakamura (1985, p. 293).

Shaw (1994) explored the dynamics underlying the dramatic rises in the labour supply of married nonblack women in the United States over a 21 year period, from 1967 to 1987, in order to study the change in persistence.⁹ She used data taken from the Panel Study of Income Dynamics (PSID) and divided her data into three different samples, married women, single women and women who have moved from being single to becoming married during the 21 year period, all split up by age groups, 25-34, 35-44, 45-54 and 55-64 years old. To investigate the change in persistence in labour supply, Shaw considered the hours of work as the labour supply variable and introduced lagged hours in the hours equation. After having controlled for other factors, such as the age and number of children and husband’s health status, her overall finding was a statistically significant persistence. Furthermore, whereas the participation rate of women has risen dramatically over the 20-year period under consideration, the degree of persistence was found to have changed a little, implying that the percentage of persistent workers¹⁰ and persistent non-workers remained almost the same in 1987 as it was in 1967. Shaw suggested that this little change in persistence is due to the fact that whereas the number of women who are persistent non-workers has decreased, they have been replaced by somewhat greater increases in the number of persistent workers. In the early 1970s, only about 28 percent of all women were persistent workers, while by the late 1980s, about 51 percent of all married women had become persistent workers. One reason for the growing number of married

⁹ Shaw (1994) refers to state dependence and persistent individual heterogeneity as “period-to-period persistence” and “lifestyle persistence” respectively.

¹⁰ Shaw (1994) classified a “persistent worker” as an individual who works four out of four continuous years.

working women is the increase in the persistence of employment from the single to the married state, which indicates that far fewer women abandon their work upon marriage, and then they maintain these working habits during their career. Similarly, young single women have become much more persistent workers, a trend that lasts longer due to marriage at later ages, and a trend that persists into their married years despite childbirth and rearing. Finally, Shaw also found that unobserved individual heterogeneity was a key element of the persistence. However, whereas Shaw (1994) pointed out the importance of distinguishing between state dependence and serial correlation, her study did not investigate whether the persistence could have resulted from unobserved transitory shocks that might be serially correlated.

Hyslop (1999), using panel data from the PSID from 1979 to 1985, estimated probit models of married women's labour force participation decisions for whom the husband was a labour force participant for all sample years. Hyslop, using maximum simulated likelihood estimation, augmented Heckman's approach (1981b) and introduced first-order serial correlation in the error term. Hence, his life-cycle model investigated all three elements, state dependence, unobserved individual heterogeneity and serial correlation. He also adopted a correlated random effects specification for which he assumes that unobserved heterogeneity is correlated with only the time varying regressors. In all his approximations, Hyslop (1999) found significant state dependence, unobserved heterogeneity, and negative serial correlation in the error term. More specifically, Hyslop obtained two main findings: (1) the first-order serial correlation in the error term is important in the model and the importance of random effects is significantly reduced when serial correlation in the error term is introduced; and (2) once serial correlation in the error term is included, the hypothesis that fertility and husband's income are uncorrelated with the random effects cannot be rejected.

Chay and Hyslop (2000) employed dynamic discrete choice models to examine both married women's labour force participation and welfare participation behaviour of women with the use of panel data from the Survey of Income and Program Participation (SIPP) and the Panel Study of Income Dynamics (PSID) in the United States. Following Heckman's (1981b) approach to the initial conditions problem, they examined both random effects and fixed effects approaches to their models and found that the reduced-form approximation to the initial conditions that allows for autoregressive error terms

gives evidence that less than forty percent of the overall persistence in married women's labour force participation is attributable to structural state dependence. The models did not include health-related variables.

Michaud and Tatsiramos (2005), similarly to Hyslop (1999), used eight waves of the European Community Household Survey (ECHP) and examined the intertemporal labour supply decisions of married women in a country-by-country analysis in Europe (France, Italy, Spain, Germany, Netherlands, and the UK) in the period 1994-2001,¹¹ in an attempt to pinpoint the differences in the labour force participation rates across countries. Presence of high persistence was found for women in all 6 countries and negatively associated with the rate of employment. Women in countries with low employment rate (the Netherlands, Italy, and Spain) experienced higher persistence than women in countries with high employment rates (France, Germany, and the UK). In particular, about 75% of women in Italy did not have any transition in their employment status within the sample period of 8 years, whereas the percentage of women who did not change employment status in the UK was about 58%. The employment rate for Italy and the UK was respectively 47.4% and 67.1 %. Moreover, in all countries, higher education was found to be associated with higher employment probability. Very good health, measured by a binary self-assessed health variable, was only significant for the UK and the Netherlands (the estimates of the dynamic probit are 0.044 and 0.046 respectively). Michaud and Tatsiramos (2005) performed Heckman's (1981b) approach and found strong presence of state dependence within each country. The average partial effects revealed that the Netherlands had the highest difference in probabilities for women who worked in the last year compared to those who did not, and Spain the lowest difference. The estimate of state dependence for the Netherlands and Spain was 0.489 and 0.276 respectively. For the UK, the estimate of the lagged employment status variable shows that women who worked last year were more likely to work in the current year by 31.1% compared to those who did not work last year. The decomposition approach of Michaud and Tatsiramos (2005) showed that differences across countries in state dependence does not explain the sample cross-country differences in employment rates, however. Rather, they concluded, the cross-country differences in employment rates are mostly due to unobservable heterogeneity and

¹¹ The sample consisted of women of working age, between 18 and 65 years old.

observable heterogeneity as captured by education, husband's income, and fertility. In fact, the segmentation of the labour market with respect to women with different education levels was found to be the most important factor.

In an extension of this work, Michaud and Tatsiramos (2008) examined again the intertemporal employment decisions of married women across 7 countries in Europe (Denmark, France, Germany, Italy, the Netherlands, Spain, and the U.K), using data from ECHP,¹² with emphasis on the fertility effects on employment. They performed their analysis with the use of dynamic probit models for which they employed Heckman's (1981b) approach to deal with the initial conditions problem, and incorporated serial correlation in the error terms similarly to Hyslop (1999) in order to examine potential bias due to the omission of serial correlation. Further, to account for the potential endogeneity of fertility when modelling women's labour supply, Michaud and Tatsiramos (2008) specified a dynamic bivariate probit panel model of employment and fertility decisions and for instrument they used the presence of children of the same sex in families with two or more children. Under the exogeneity assumption of employment decisions, Michaud and Tatsiramos (2008) found evidence of positive state dependence for all countries, after controlling for observed and unobserved individual heterogeneity, and serially correlated unobserved transitory shocks. In the UK, their estimate of state dependence is 0.582 and significant at the 1% level, which implies that the probability of a woman being employed is 58.2 percentage points higher if she was employed than if she was non-employed in the previous year. The estimated autocorrelation coefficient is -0.269 for the UK and the estimate for health (measured by a binary self-assessed health variable) exhibited significant negative effect only for the UK and the Netherlands, at the 1% significance level (-0.062 and -0.109 in the two countries, respectively). Under the endogeneity assumption in which employment and birth equations are jointly estimated, all countries exhibited significant positive state dependence, but smaller compared to the ones under the exogeneity assumption. In the UK, the estimate of state dependence becomes 0.452. Regarding the coefficient of health Regarding the coefficient of health, the UK obtains a significant estimate of -0.049 which is smaller compared to the estimate under the

¹² The unbalanced sample consisted of married or cohabitating women aged between 20 and 45 years old for whom information was observed for at least three waves from 1994 to 2001. Employed women were considered those working more than 15 hours per week.

exogeneity assumption. Their analysis provided clear evidence that ignoring both endogeneity of fertility and serial correlation causes the coefficients of interest to be biased upwards.

Keane and Sauer (2009), using maximum simulated likelihood, estimated a generalized autoregressive dynamic model in which they incorporated classification error in reported employment status.¹³ Keane and Sauer (2009), using the same data with Hyslop (1999), the same approximate solution for the initial conditions as in Heckman (1981b), and the same specification for correlated random effects as in Hyslop (1999), concluded that their generalization, ignoring classification error, slightly decreases the effect of state dependence in employment from 1.042 to 1.031. In contrast, their generalization increases the variance of the individual effect from 0.485 to 0.519 and the first-order autocorrelation coefficient from -0.213 to -0.141. Their generalization results in a minor improvement of the log-likelihood by 2 points.¹⁴

Edon and Kamionka (2010) jointly modeled employment and fertility decisions of women in France, Spain, Germany, UK, and Denmark using the first eight waves from the ECHP (European Community Household Panel), from 1994 to 2001.¹⁵ Their purpose was to investigate the determinants of female labour supply and the feedback effect of participation decisions on fertility decisions, given the different institutional and social environments of these five countries. Using as instruments an indicator of whether the first two children have the same gender and another indicator of whether the oldest child is a boy,¹⁶ they estimated a dynamic bivariate probit model with random effects which takes into account the initial conditions. Their empirical findings showed evidence of significant true state dependence in participation decisions, most likely because of the institutional or social environments. Their findings also demonstrated large individual unobserved effects

¹³ Classification error happens in micro data sets when the employed individuals falsely report being unemployed and the unemployed individuals falsely report being employed (Keane and Sauer, 2009, p. 976)

¹⁴ Results mentioned are in first two columns of table IV, page 985 of Keane and Sauer (2009) for which classification error is not accounted for.

¹⁵ The sample consisted of women aged between 20 and 56 years old with average age varying between 42 and 44 years old and which were continuously living in couple (married or not) during the sample period. The participation decision is indicated by a binary variable whether a woman participates in the labour market works, and the fertility decision by a binary variable whether a child was born during the last twelve months.

¹⁶ Edon and Kamionka (2010) consider these two instruments exogenous to participation decision assuming that they are independent of the error term in the participation equation, but highly correlated with the fertility decision.

("spurious state-dependence"), most likely because of strong preferences of women for maternity which lead them to weaker preference for participation in the labour market. For the labour supply decision, in particular, the findings are in line with previous research. In all countries, education and transitory income have a positive impact on employment and permanent income has a negative impact. For the effect of children, Edon and Kamionka distinguished it into a negative direct effect, which is higher in countries where the duration of maternity leave is longer (for Denmark and France), and a dynamic effect whose intensity depends on the institutional environment, such as childcare availabilities (for Denmark and France), labour market flexibility (for Denmark and UK), and cultural parameters such as family ties (for Spain). In the UK, the effect of the number of young children, aged between 1 and 3, is stronger than the current effect of a birth. According to the authors, the labour market is flexible in the UK, whereas childcare provision is limited and benefits are means-tested.¹⁷ Thus, as the authors argue, this effect of the number of young children, aged between 1 and 3, indicates that, frequently, women do not stop working for childbearing in the UK because the maternity leave is unpaid. However, mothers often leave the labour market during the first year after the birth of a child, especially if they have other young children (Edon and Kamionka, 2010).

In France, Collet and Legros (2016), using data from the French Labour Force surveys from 1997 to 2002, constructed an intertemporal model of labour force participation for married women and estimated a dynamic probit model with correlated random effects that can accounts for state dependence, initial conditions, and unobserved heterogeneity. Their empirical results highlighted the crucial role of all three components in women's decisions to work and negative serial correlation in the transitory error component. The presence of young children reduces female labour participation, whereas non-labour income increases female labour participation initially, but decreases it afterwards.

¹⁷ A means-tested benefit is a payment available to individuals that can prove that their income and capital (their 'means') are below defined limits. In the United Kingdom, means-tested benefit is a core feature of the welfare state.

2.3.2 Structural approximations in female labour supply studies

We do not attempt to provide a full review of the structural dynamic labour supply models. Nonetheless, we provide a short review with the research of Eckstein and Wolpin (1989) as our benchmark, because their research is the first to adopt a full solution approach to modelling female labour supply. Eckstein and Wolpin (1989) solved the female utility maximization problem with a structural model, which means that the parameters to be estimated are derived from objective functions and constraints, founded in the economic theory. They solve a dynamic programming model whereby a woman maximizes the present value of her utility by choosing whether to work or not, subject to her budget constraint. In other words, her decision rule for whether to work or not is simply to work if the wage offer exceeds the reservation wage.¹⁸ Specifically, the female wage is considered to be a random variable that depends on work experience and the dynamics arise because the woman chooses whether to work or not at time t because of the effect of previous accumulated experience on her wage. The estimation of the dynamic model is achieved by maximum likelihood methods.

There have been many extensions of Eckstein and Wolpin's (1989) work. An important extension, among others, is that of Van der Klauuw in 1996. Van der Klauuw (1996) constructed a similar dynamic model of female labour supply, where the marital status is considered as an endogenous decision that interacts with labour supply decisions. A difference from the Eckstein and Wolpin (1989) model is that Van der Klauuw (1996) represents the birth of children by an exogenous stochastic process, conditional on the female marital status. Van der Klauuw (1996) modeled marital status decisions and labour supply decisions for women who have left school, which may be as young as 14 years old. A last but distinct difference is that Van der Klauuw (1996) did not assume a household utility function. Instead, a woman receives utility from her own income and only a fraction of her husband's income, in case she is married. Altug and Miller (1998) extended similarly Eckstein and Wolpin's (1989) work. Hours worked and hours spent out of the work are included in their dynamic framework, whereas Eckstein and Wolpin (1989) only explore binary discrete choices for work. A second difference is that Eckstein and Wolpin

¹⁸ To avoid having to model fertility decisions, Eckstein and Wolpin focused only on women who were at least 39 years old in 1967 and hence beyond fertility age.

(1989) assume that neither saving nor borrowing occurs, whereas Altug and Miller (1998) assume that all resources are allocated optimally. Thus, their dynamic framework accounts for the effects of aggregate shocks on labour supply decisions and, therefore, contains idiosyncratic shocks in preferences that affect the utility associated with different consumption and labour supply decisions at the individual level.

Francesconi (2002) extended Eckstein and Wolpin (1989) by the evaluation of fertility as a decision, identified only by the number of children and not by their age, and by the inclusion of full-time and part-time work in separate wage functions. In 2002, Keane and Wolpin estimated a structural dynamic model in which women make sequential joint decisions on full-time and part-time work, marital status, fertility, education, and welfare participation. In their model, in contrast to Francesconi (2002), fertility decisions depend on both the number of children and their age and a time cost of rearing, in accordance with the current age distribution of children. The important feature of Keane and Wolpin's (2002) work is that it offers a link between the literature on dynamic labour supply and the literature on the welfare. The participation of women in public welfare programs is included as a decision. Lagged welfare participation affects labour market supply and marital status decision. More recently, Keane (2010) augmented the dynamic framework by investigating possible discrimination against minorities. The focus of the article is on the possible differences between race groups in labour supply, marital status, fertility, education and welfare participation decisions, and the impact of changing welfare rules.

Similarly to the above approaches, Ge (2011) specified a dynamic framework where women jointly decide each year from the age of 22, that is, after their graduation from high school (and possibly from a higher education institution), upon labour supply, their further education and marital status. Eckstein and Lifshitz (2011) modified Eckstein and Wolpin (1989) and set the first period of optimization at age 23. At that age, women are supposed to have completed their education. However, fertility is not treated as given. Eckstein and Lifshitz (2011), similarly to Eckstein and Wolpin (1989), kept work experience as the only endogenous variable, whereas they formed transition functions for the probability of having another child (as in Van der Klauuw, 1996), the probability of getting married and the probability of getting divorced. Eckstein and Lifshitz (2011) assumed that these demographic characteristics have expectations that are potentially

important in the formulation and estimation of a dynamic model of female labour supply model.

2.4 Health

2.4.1 Health. Definitions and measures.

A widely accepted definition of health is that given by the World Health Organization (WHO) in 1948: “a state of complete physical, mental and social well-being, not merely the absence disease or infirmity,” while Chirikos and Nestel (1981, p. 301) define health to be the “bundle of physical and mental capacities affecting the ability to perform primary and secondary social role responsibilities.”

Through both of these definitions it becomes clear that health is a multi-dimensional concept. From the first definition of health, we conclude that health is not restricted to physical (i.e., somatic) health. It also encompasses mental health as an integral part of an individual's overall health. Mental health could be described as a condition that is associated with a person's psychological and emotional well-being and is used as a term to describe how people evaluate their lives and deal with the usual stressful events, in the absence of a mental disorder. From the second definition, it also becomes clear that the definition of health depends in part on the questions the researchers address. If we examine the effect of health on labour supply, we need a measure of health to be an accurate measure of the full extent to which individuals' employment status is affected by health.

Nonetheless, “true” health is not observable. Besides, among researchers, there is no consensus about the measurements of health that are more accurate as an indicator for health which affects labour supply. In turn, researchers rely on the use of health measures found in administrative surveys. The health measures more often available in administrative surveys are subjective, self-assessed health measures and these are the most commonly used measures of health status in studies that investigate the impact of health on labour market outcomes and social and economic inequalities, such as poverty (for example, Biewen, 2009).

Among the self-assessed health measures, one of the most frequently used is a question that captures a person's perception of his or her own health at a given point in

time and provides an ordinal ranking of perceived health status. For example, the British Household Panel Survey (BHPS) asks the respondents to classify their health status according to the following question: “Please think back over the last 12 months about how your health has been. Compared to people of your own age, would you say that your health has on the whole been...” with five response categories “excellent,” “good,” “fair,” “poor,” and “very poor.” A similar question appears on the European Community Household Panel User Database (ECHP-UDB) that allows a comparison across countries. In the ECHP self-assessed health status is measured as “very good,” “good,” “fair,” “poor,” or “very poor.” Unlike the BHPS, respondents are not asked to compare themselves with others of the same age. Another self-assessed health measure is whether there are limitations on the daily activities. For example, the BHPS asks the respondents the following question: “Does your health in any way limit your daily activities compared to most people of your age?” Respondents are left to identify their own perception of health and their daily activities. Similarly, the ECHP asks respondents the question, “Do you have any chronic physical or mental health problem, illness or disability?” With response categories “yes” or “no” and if the response is “yes,” the respondents are asked a second question: “Are you hampered in your daily activities by this physical or mental health problem, illness or disability?” With response categories “no,” “yes,” “to some extent, yes” and “severely.” A third frequently used self-assessed health measure is whether there are chronic health problems. For example, the Survey of Health, Ageing and Retirement in Europe (SHARE) survey, whose samples are from the population aged 50 years and above, asks the respondents to classify their own health in six domains: mobility, cognition, pain, sleep, breathing, and emotional health. The response categories are “none,” “mild,” “moderate,” “severe,” and “extreme.” In addition, for each domain, respondents are asked to evaluate three vignettes, which describe hypothetical cases, and individuals are asked to rate them in the same way, i.e., on the same response scale, as they would evaluate their own health.

Among other self-assessed health measures are whether there are health limitations on the ability to work, whether there are chronic conditions (such as asthma, diabetes, and arthritis) and whether there are acute conditions (such as cancer, stroke, and heart attack). Among objective health measures, we could mention clinical assessment of health problems, expected or future mortality, the utilization of medical care (such as hospitalization and out-of-hospital medical services), the construction of health scores, and

the construction of indices for mental and physical conditions. The vast majority of studies using data from developed countries concentrate on self-assessed health status, health limitations or utilization of medical care (Currie and Madrian, 1999).

Despite the fact that subjective self-assessed health measures are not equivalent to objective health measures, they have been used widely in studies of the relationship between health (physical and mental) and hourly wages (Contoyannis and Rice, 2001); between health and socioeconomic status in the UK (Contoyannis et al., 2004); between health and poverty (Biewen, 2009; Andriopoulou and Tsakloglou, 2011). These measures of health have been found to be powerful predictors of early retirement decisions due to health shocks (Jones, Rice and Robert, 2010). Not only have subjective self-assessed health measures been found to efficiently measure the overall health of individuals, but they are also increasingly proved to be powerful predictors when used to assess the impact of self-assessed health on future health service utilization (doctor visits and hospital admissions) across illness groups in Australia (Doiron et al., 2015). Moreover, Doiron et al. (2015) conclude that, despite their low predictive power compared to objective health measures, subjective self-assessed health measures are a valuable indicator of objective health and could be used even in the absence of comprehensive and objective health measures.

2.4.2 Endogeneity of health in labour supply models

Most of the literature that examines the impact of health on labour outcomes treats health as an exogenous variable.¹⁹ Researchers assume, explicitly or implicitly, that the variation in health is generated by exogenous shocks to health. Chirikos (1993) concludes about the exogeneity of health in labour supply models, that, at least to some degree, “health matters” in the determination of every labour market outcome. Similarly, Currie and Madrian (1999) explain that “this may not be an unreasonable assumption given that current health depends on past decisions and on habits that may be very difficult to break, and the fact that individuals often have highly imperfect information about the health production function at the time these decisions are made” (page 3313).

¹⁹ Chirikos (1993) and Currie and Madrian (1999) have comprehensive overviews of such studies.

However, as we mentioned earlier in the introduction (page 7), a concern that emerges in the recent literature is that health might be endogenous in labour supply models. Grossman (1972) gave the benchmark to consider that health, while it is considered to be predetermined as an “endowment” at birth, it depreciates over the course of life. Therefore, individuals must constantly replenish the initial “stock” of health if they want to reduce the depreciation of their health. They can presumably achieve this by investing in their health by sacrificing time²⁰ and monetary resources. For example, individuals might need to spend more time for relaxation, recreation or exercise and more income to cope with increasing medical expenses. In turn, the availability of time and monetary resources may depend on the individual’s labour supply, past and current. This implies that individuals decide upon their labour supply simultaneously with health, as they would do with other commodities. Therefore, Grossman’s view leads researchers to treat health as endogenous.

We also mentioned earlier in the introduction (page 8) that, beyond the considerations based on Grossman (1972), a second empirical source of the endogeneity of health is that the relationship between health and hours worked can also work in the opposite direction whereby employment can affect directly health (Currie and Madrian, 1999). For example, working in a stressful environment or being in unemployment for a long period may have detrimental consequences for mental health (Clark and Oswald, 1994). Or there might be hazardous working conditions that may lead to deterioration of health, such as manual jobs with high physical demands that have a high risk of injury. On the other hand, good social interactions at work, personal fulfillment from employment engagement, and the self-esteem that employment brings may benefit mental health (Cai and Kalb, 2006). Thus, employment status could also affect health, although the direction of the impact is ambiguous. But, regardless of the empirical validity for the dual causality between employment status and health, the possibility that labour supply and health influence each other simultaneously further suggests that health should be treated as endogenous (Currie and Madrian, 1999).

In addition, another source of endogeneity of health may be unobserved individual heterogeneity. This endogeneity bias derives from the presence of unobserved individual

²⁰ Grossman (1972) holds the view that health capital varies from other types of human capital. Health, unlike education, does not affect individuals’ productivity but affects the amount of “healthy time” available to the individual.

characteristics which may affect participation status, but cannot be explicitly included in the model because they are not observed. For example, both employment status and health may be jointly affected by a third, unobserved, variable such as individual preferences, ability, and taste shifters. If unobserved individual heterogeneity is ignored in the estimation, the coefficient of health is likely to be overestimated. The endogeneity, derived from all these three sources, is called “true” endogeneity bias.

A fourth source of endogeneity is likely to be heterogeneity, caused in particular by the use of self-assessed health measures, and results in “reporting” bias. Self-assessed health measures are subjective proxies of “true” health. Therefore, it is almost certain that they are prone to measurement error. The error may be due to the administration effect (for example, self-completion questions reveal more morbidity than face-to-face interviews), to the framing (or learning) effect (for example, respondents are influenced by the morbidity questions), or due to varying answers influenced by individual characteristics. For example, Hernández-Quevedo, Jones, and Rice (2005) argue that “the systematic use of different threshold levels by sub-groups of a population reflects the existence of “reporting” bias. These differences may be influenced by, among other things, age, gender, education, income, language, and personal experience of illness. This means that different groups appear to interpret the question within their own specific context and, therefore, use different reference points when they are responding to the same question” (page 6).

Last, measurement error may also be due to the fact that respondents who choose not to work for non-health reasons rationalize their decision by reporting poor health. This is the case in which an unemployed individual may seek “socially accepted” justification for being out of the labour market or “confirm” eligibility for a social security benefit. In the latter case, the measurement error is not random. Consequently, bias arises when self-assessed health status is used to ‘justify’ a prior work decision. This is named “justification” bias (also referred to as “rationalization” bias). In turn, when there is justification bias, the effect of health is likely to be overestimated. The majority of studies that investigate “justification” bias focus explicitly on samples of older individuals and on the retirement-health nexus rather than on younger individuals.

Therefore, despite their popularity and their ability to encompass a great number of health conditions, self-assessed health measures are of questionable reliability. More subjective measures are less prone to measurement error but, as Bound (1991) argues,

neither should they be considered exogenous. Moreover, Hernández-Quevedo, Jones and Rice (2005) state that “objective measures of health status are rare in survey data and where they do exist they are often too specific to particular health conditions. Accordingly, their applicability as an overall measure of an individual’s health status is often limited” (page 23).

Attempts to surmount heterogeneity would be to estimate a self-assessed, and potentially endogenous, health measure as a function of more ‘objective’ health measures and together they could then be used to define a latent ‘health stock’ variable, purged from heterogeneity. The latent ‘health stock’ variable would then be used in the labour supply model. A representative work of this kind of research is that of Lindeboom and Kerkhofs (2009). Using data from the first two waves of the Leiden University Center for Research on Retirement and Aging (CERRA) panel survey in the Netherlands, these authors explored the interrelation between health and transitions from work to unemployment, disability or (early) retirement of older workers. The self-assessed health was estimated as a function of more ‘objective’ measures of health and demographic characteristics in order to construct a latent ‘health stock’ variable with thresholds dependent on the labour market status.²¹ Thus, the effect of labour supply on the thresholds could be interpreted as “reporting” bias, in the sense that reporting of health may be influenced by participation status. This health stock variable, was then used as a proxy for health in their model of retirement. The empirical results showed evidence of substantial “justification” bias and endogeneity of self-assessed health.

In this framework, but with results in favour of absence of “reporting” bias of the self-assessed health measures, Doiron, Fiebig, Joharb, and Suziedelyte (2015) combined data derived from two databases, the “45 and Up Study”²² and the Medicare database in

²¹ The subjective self-assessed health measure in their analysis was the response to the question 'Does your health limit you in the kind and the amount of work that you can do?' grouped in four categories; 'causes no problem', 'causes some problems', 'causes severe difficulties' and 'makes it impossible to work'. The more objective measure was a constructed total health score derived from responses to 57 health-related (somatic and mental) questions of the Hopkins symptoms Checklist whose index ranged from 0 to 171 and its increase represented changes from best health to worst health.

²² The “45 and Up Study” is a cross-section survey with health administrative data for individuals aged 45 and over in the state of New South Wales and the Medicare database is administrative individual medical records that covers all permanent residents in Australia since, essentially, all permanent residents have access to public health insurance.

Australia to explore the predictive power of self-assessed health over an extensive set of future health outcomes. The health outcomes were measured by hospitalizations, out-of-hospital medical services and prescription drugs. To assess the predictive power of self-assessed health, they compared self-assessed health to two groups of more objective health measures. The first group comes from the “45 and Up Study” and consists of 15 variables related to self-reported diagnoses of illnesses, such as cancer, heart disease, stroke, diabetes, asthma, Parkinson’s disease, depression and anxiety, and daily health limitations that consist of physical health limitations and mental distress, as measured by the Kessler psychological distress scale. The second group of objective health measures includes in total 136 variables related to past illness-specific hospital admissions in the past five years and out-patient service and prescription drug use in the last 12 months from the survey date. Doiron et al. (2015) found evidence of significant impact of self-assessed health on future health service utilization across all illness groups. Specifically, they found that poor self-assessed health substantially increases the utilization of health care services, especially specialist visits and hospital admissions, with the impact to be larger for females and the youngest. However, compared to health administrative data and objective health measures, self-assessed health has less predictive power. Specifically, objective health measures derived from past administrative data, such as diagnoses from past hospitalization, were found to be the most predictive of an individual’s future utilization of health care services, followed by survey-based self-reported health measures. Interestingly, however, self-assessed health was found to predict serious and chronic illnesses, such as cancer, better than less serious illnesses. Doiron et al., therefore, concluded that, despite its low predictive power compared to objective health measures, self-assessed health is a valuable indicator of objective health and could be used even in the absence of comprehensive and objective health measures.

In the UK, Hernández-Quevedo, Jones, and Rice (2005) explored reliability in the measure of self-assessed health and investigated the extent to which the different wording to the self-assessed health variable in the ninth wave of the BHPS, contributes to “reporting” bias or equally, as stated by the authors, “contamination by measurement error.” In other words, the attention was whether different groups of individuals, defined through socio-economic characteristics, respond in different ways to the change in the measurement instrument. Specifically, for waves 1-8 and 10 onwards, the self-assessed health variable in the BHPS corresponds to “health status over the last 12 months” and

respondents are asked: “Compared to people of your own age, would you say your health over the last 12 months on the whole has been: excellent, good, fair, poor, very poor?”. However, in wave 9, the self-assessed health variable changed since then and it represents “general state of health” in which the respondents are asked: “In general, would you say your health is: excellent, very good, good, fair, poor?” Hernández-Quevedo, Jones, and Rice (2005) explored three distinguishable differences between the two wordings of the question. In wave 9, the self-assessed health question does not include the age benchmark. This means that individuals are not asked to assess their level of health “compared to people of your own age” and the question does not include the time frame of reference, “over the last 12 months.” The third difference is a modification to the response categories. Although both questions provide five possible answers to the respondents, the response category “very poor” is not available in wave 9, but “very good” is added between “good” and “excellent.” Hernández-Quevedo, Jones, and Rice (2005) found a statistically significant index shift at wave 9, which means that the distribution of self-reported health was different at wave 9 than at waves 1–8 or waves 10–11, indicating an index (parallel) shift. Their findings, however, indicated that the extent of this parallel shift does not vary by socio-economic characteristics. Finally, they concluded that collapsing the categories of self-assessed health over the two versions of the question is best achieved by collapsing to a four-category version of self-assessed health by combining “very poor” and “poor” health in waves 1-8 and 10-11 and “good” and “very good” in wave 9. This methodology was found to not affect the estimated relationship between self-assessed health and socio-economic characteristics and is widely adopted by many authors who use BHPS data in studies in which self-assessed health is of prime concern. Interesting examples are the research by Brown, Roberts, and Taylor (2010), who used 14 waves of the BHPS to explore the effect of health on stated reservation wages of unemployed men and on market wages of employed men, and by Jones et al., (2010) who used 12 waves of the BHPS in an attempt to investigate potential reporting bias in self-assessed health and self-assessed health limitations from individuals in order to justify early retirement decisions due to health shocks.

Heterogeneity (and “reporting” bias) in self-assessed health has also been identified by Pfarr, Schmid, and Schneider (2012) in a cross-country comparison of eleven

countries.²³ The authors argued that observed and unobserved heterogeneity play an important role when analyzing self-assessed health variables in a cross-country comparison and both have to be taken into consideration when modelling self-assessed measures with a generalized ordered probit model.

2.4.3 Links between labour supply and health

Cai and Kalb (2006) examined the effect of health on labour force participation in Australia using the first wave of the Household, Income and Labour Dynamics (HILDA) Survey, conducted in 2001. They split their sample into four groups by sex and two age categories, men and women aged less than 49 and men and women aged 50 or over, and attempted to explore potential endogeneity of health to labour force participation, by estimating a health equation and a labour force participation equation simultaneously. The full information maximum likelihood estimation method was used to estimate the model. In the health equation, they included a dummy variable for the presence of long-term self-assessed health conditions, a dummy variable for lack of physical activity, and a summary indicator for specific physical functioning limitations (such as climbing stairs), all treated as exogenous because they are unlikely to be influenced by current labour force participation. The null-hypothesis of exogeneity of health to labour force participation was rejected for all four groups. Overall, Cai and Kalb (2006) found that poorer health decreases the probability of labour force participation for all four groups and that the effect is larger for the older groups and for women especially. Moreover, their results indicated that women have lower labour force participation rates than men for all health categories and the older women have, additionally, lower labour force participation rates than the younger women. Specifically, for a woman aged 50 or over, deterioration of health from excellent to poor reduces the probability of labour force participation by 25.32%. For a younger woman aged below 50, the same change in health reduces the probability of labour force participation by less than 7%.

²³ The sample consists of individuals in Austria, Germany, Sweden, the Netherlands, Spain, Italy, France, Denmark, Greece, Switzerland and Belgium. The sample included men and women aged above 50 years old.

Haan and Myck (2009), using twelve years of data from the German Socio-Economic Panel study²⁴ investigated the association between health and employment and, extending the approach set out in Alessie, Hochguertel, and van Soest, (2004), employed a joint intertemporal model of health and non-employment in which they assumed that the health in the previous year affects employment in the current year and the employment in the previous year affects health in the current year. Thus, they estimated a dynamic bivariate logit model that accounts for persistence effects, the joint distribution of unobserved heterogeneity and controls for the initial conditions as in Wooldridge (2005). The health indicator is based on the survey question asking respondents to rate their general health on a five-point scale; “very good,” “good,” “satisfactory,” “poor,” “very poor.” The health-related variable is then constructed as a binary variable with those reporting either “poor” or “very poor” health classified as being in “poor health”. “Non-employed” are those who report zero working hours during the last week of the interview. The empirical results highlighted the significant role of persistence in the dynamics of poor health and non-employment, unobserved heterogeneity and the initial conditions. Haan and Myck (2009) showed that poor health in the previous year is a significant determinant of non-employment in the current year and non-employment in the previous year has a positive effect on the probability of being in poor health in the current year. They also demonstrated the significant role of ageing in determining the two outcomes, that is, both effects increase with age. Finally, Haan and Myck (2009) highlighted that ignoring unobserved heterogeneity in the estimation leads to an overestimation of the relationship between health and non-employment. Ignoring unobserved heterogeneity, the estimates of the coefficients on the lagged endogenous variables (poor health and non-employment) become upwards biased and the persistence of health and non-employment (of both processes/outcomes) augments. Still, after controlling for the correlated unobserved heterogeneity which influences both processes, Haan and Myck (2009) found that the effects of health on employment are significant.

In 2010, Cai (2010) repeated the same estimation procedure and methods as in Cai and Kalb (2006) by employing data from the first four waves of the HILDA Survey. The panel data nature of the data permitted him to control for potential individual heterogeneity

²⁴ The sample consisted only of men aged 30-59 for a sample period from 1996 to 2007.

in his attempt to better assess the relationship between health and labour force participation. The sample is divided into two groups, men and women between 25 and 64 years. Similarly to his previous research (2006), Cai (2010) provided evidence that health has a significant effect on labour force participation for both groups. The worse the self-assessed health is, the lower the labour force participation rate is. Especially for women, a change from excellent to poor health reduces the probability of participation by 13 percentage points. In addition, the null-hypothesis of exogeneity of health to labour force participation was rejected for both men and women.

Oguzoglu (2010), using a sample of men aged 24 to 64 years old and of women aged 24 to 60 years from the first five waves of the HILDA Survey, examined whether or not work limitations²⁵ impact work decisions (to be a labour force participant or not) with dynamic panel data models in which state dependence, endogenous initial conditions and unobserved individual heterogeneity are accounted for. Oguzoglu (2010) employed, along with a one-equation for labour force participation, a joint estimation of the equation for labour force participation with a work disability reporting equation and modelled the correlation between unobserved components in the two-equation setup because unobserved individual characteristics that make individuals more likely to be out of the labour force could also make them more likely to report a work limitation. This, in turn, could make work limitations endogenous in the labour force participation equation. For this, he also modelled the correlation between unobserved heterogeneity and exogenous variables as in Mundlak's (1978) approach. For women, the findings verified that a significant correlation exists between unobserved components of labour force participation and work disability equations and showed that ignoring this correlation can markedly overestimate the impact of work limitations on labour force participation in the one-equation setup. Except for the effect of persistence due to individual heterogeneity, the results also verified the significant effect of state dependence; both men and women who were in the labour force at time $t - 1$ are more likely to be in the labour force at time t , than individuals who were not in the labour force at time $t - 1$.

²⁵ The respondents were asked the question: "Does your condition limit the type of work or the amount of work you can do?" A binary indicator for the presence of work limitations is constructed from this question.

Oguzoglu (2016), using a sample of men and women aged 24 to 65 years from the first seven waves from the HILDA Survey, employed a dynamic mixed multinomial logit model that allows for state dependence, endogenous initial conditions and unobserved individual heterogeneity, to examine whether or not a work limiting disability²⁶ affects individual decisions to leave or stay in an employment state.²⁷ The findings demonstrated a highly significant effect of disability on employment outcomes. For example, women with work limitations are 8.4 percent less likely to be full-time workers, compared with those without disabilities. The average partial effects also demonstrate these findings; women who become disabled at time period t and were full-time employed at time period $t - 1$ are more likely to transit into part-time work by 4.8 percent, than their non-disabled counterparts. The results also highlight the great degree of state dependence in all employment outcomes for women. Full-time working women are, on average, 65 percent more likely to be working full time in the next period compared to women who are not currently participating in the labour force and part-time working women are, on average, 41 percent more likely to be working part time, respectively.

Recently, Cai (2018) explored the effect of various observed factors, such as education, age, health and the number of children of different ages on the extensive and intensive margins of the labour supply of married Australian women with the use of the first 13 waves of the Household, Income and Labour Dynamics in Australia Survey.²⁸ Moreover, he examined potential persistence in labour supply and its sources. He employed a dynamic²⁹ two-tiered tobit model which he estimated using the maximum simulated likelihood estimator (MSLE). The estimated marginal effect for the state dependence is 0.482, suggesting that a woman is more likely by 48 percentage points to work in the current year if she worked in the previous year, compared to a married woman who did not work in the previous year. The estimate for the working hours conditional on

²⁶ The respondents were asked the question: “[...] do you have any long-term health condition, impairment or disability that restricts you in your everyday activities, and has lasted or is likely to last, for 6 months or more?”

²⁷ The employment states were categorized into four groups; full-time employed, part-time employed, unemployed and not in the labour force.

²⁸ The sample consisted of married women of working age between 20 and 60 years old.

²⁹ Lagged labour supply enters the equation as two variables – lagged employment status and lagged hours worked conditional on being employed.

being employed was 0.59, implying that an increase of 1 hour per week in the previous year raises working hours in the current year by about 0.6 hour per week if a woman works in both years. Subsequently, both employment and hours worked exhibited positive state-dependence for married women. Cai (2018) also found evidence of presence of unobserved heterogeneity and negative serially correlated unobserved transitory shocks as well as evidence that state dependence remains even after controlling for these sources of persistence. The only measure of health used in this analysis was a binary variable indicating whether an individual has a long-term health condition. Cai (2018), finally, found women with a long-term health condition are less likely to work, compared to women without long-term health conditions and that young children under the school age, non-labour income and education are dominant determinants of women's labour supply.

Very recently, Delattre, Moussa and Sabatier (2019), using data from French panel survey, used the Granger causality framework and employed a dynamic bivariate probit model that accounts for persistence effects, the initial conditions problem,³⁰ and unobserved heterogeneity in order to test for Granger causality (i.e., predictability) between health and employment status. The health indicators are a self-assessed measure of whether individuals have encountered illness during a given year, a constructed indicator of illness severity which index measures the risk of death, a constructed indicator of disability which index measures the impact of health on individual's daily life and a percentage measure that indicates the extent to which severity or disability take place during the entire working life of the respondent. Their results indicate persistence both in employment status and health status. Hence, respondents who reported an illness in the previous year are more likely to report an illness in the current year and respondents who were unemployed in the previous year are more likely to be unemployed or out of the labour market in the current year. Moreover, respondents who reported illness in the previous year are more likely to be unemployed in the current year and respondents unemployed in the previous year are more likely to report illness in the current year.³¹

³⁰ Delattre, Moussa and Sabatier (2019) followed Wooldridge (2005) for an alternative conditional maximum-likelihood estimator for the case of serially independent errors.

³¹ The authors argue that two factors could explain these results: 1) It could highlight a job quality effect. If employment takes place under poor conditions, employment could raise the probability of illness. 2) In France, the health-care and insurance system is generous for employed individuals. They may, for example, schedule regular appointments with their doctor, giving them access to more effective health monitoring.

Roos, Lahelma, Saastamoinen, and Elstad (2005) used cross-sectional surveys in the Nordic countries³² and employed multivariate logistic regression analysis to investigate potential association between employment status and health among men and women. Furthermore, they explored the extent to which marital status and parental status contributed to that association. Two measures of health were included in the analysis, the self-assessed health grouped into two categories: “perceived health as below good” and whether the individual has limiting longstanding illness.³³ Adjusted for age and education, their findings showed evidence of negative association between employment and poor health captured by “perceived health as below good” and also negative association between employment and limiting longstanding illness. In all four Nordic countries, unemployed women were more likely to report perceived health as below good and more likely to report limiting longstanding illness than the employed women. Marital status and parental status estimates showed a modest or no impact on the association between employment status and health, with the association being slightly strengthened among women in Denmark and Sweden. More specifically, the results indicated that full-time employment was associated with better health among women with a spouse and children in a similar way to that of women without a spouse and children. The authors (Roos et al., 2005) justified this pattern strengthening the importance of social and labour market characteristics in the Nordic countries (for example, social policies and equal opportunities policies).

Consequently, they may be more likely to detect and report a disease (Delattre, Moussa and Sabatier (2019, p. 8).

³² The samples consisted of data drawn from national surveys in Denmark, Finland, Norway, and Sweden, for men and women aged between 25 and 49 years old, for sample years 1994 (for Finland and Denmark) and 1995 (for Norway and Sweden). Employment status for women was divided into 4 categories (full-time employed, unemployed, part-time employed and housewives).

³³ Individuals were asked a question: “Do you have any longstanding illness, disability or infirmity?” If the response was “yes”, individuals were asked a second question: “Does your illness/disability restrict your work or does it limit your daily activities [gainful employment, housework, schooling, studying?” Those who replied that their illness restricted their activities, at least to some extent, were classified as having “limiting longstanding illness”

2.4.4 UK studies on the links between labour supply and health

Self-assessed health measures from the BHPS are common measures used in the research of the effect of health on labour market outcomes. Arulampalam, Booth, and Taylor (2000) estimated dynamic binary panel data models to explore dynamics in unemployment incidence related to demographic characteristics and health. The sample was drawn from the first 5 waves of the BHPS and consisted only of men aged 16 or above and health was measured by a binary indicator whether there is health condition that limits the type and/or amount of work. Their results suggested that, after controlling for observable characteristics, there is strong state dependence in unemployment with respect to previous unemployment.³⁴ In fact, men aged 45 or over are more likely to be unemployed and the presence of health condition, along with age and qualifications, raises the probability of unemployment. The authors argued that state dependence in unemployment may be due to human capital depreciation, or because employers use the prior labour market experience of an employee as an indication of productivity, or because unemployed individuals are more likely to work in low-quality jobs that are marked by high rates of job loss.

It is also worth mentioning Contoyannis et al. (2004) who used self-assessed health using data from the BHPS in order to examine the effect of income, among other factors, on health. Specifically, Contoyannis et al. (2004), using data from the first eight waves (1991–1998) of the BHPS, investigated the potential persistence of health as captured by estimates of state dependence and individual heterogeneity. The analysis was conducted with the use of dynamic panel ordered probit models on samples, split by both gender and the highest academic qualifications attained at the beginning of the survey. The findings showed that the effect of individual heterogeneity is significantly reduced by the incorporation of lagged income in the analysis, but remains high and accounts for 30% of the unexplained variation in health in all model specifications. In addition, after controlling for state dependence and individual heterogeneity, Contoyannis et al. (2004) found that there is a clear association between educational attainment and self-assessed health for

³⁴ Arulampalam, Booth and Taylor (2000) refer to the effect of state dependence in unemployment as the “scarring” effect of unemployment. They define this as the implications of the past unemployment history of an individual for his subsequent experience in the job market.

women but not for men. The permanent income effect is also significant for self-assessed health. Current income, which may capture transitory income shocks, is not significant.

Hernández-Quevedo, Jones, and Rice (2005) used the first eleven waves of the BHPS in order to explore the association between self-assessed health and marital status, ethnicity, years of education, size of the household, number of children in the household, income (expressed as the logarithm of real income, adjusted using the Retail Price Index and equivalised by the McClement's scale to adjust for household size and composition) and a quartic polynomial function of age.³⁵ The expected positive relationship between income and good health was confirmed, for both men and women.

Stewart (2007) examined the extent of state dependence in unemployment and its dynamic interaction with low-wage employment, using data from the first six waves (1991–1996) of the BHPS.³⁶ A health indicator included among the demographic characteristics is whether health limits type or amount of work. Stewart (2007) estimated dynamic random and fixed-effects models separately for unemployment and low-wage employment, and also a bivariate model to explore potential endogeneity in the two states, unemployment and low-wage employment. Stewart (2007) confirmed unobserved heterogeneity that exhibits persistence over time and highlighted its importance in the results. The findings indicated that ignoring unobserved heterogeneity leads to an overstatement of the true state dependence in unemployment and if taking unobserved heterogeneity into account, this reduces the state dependence by about a third. Regardless of this, the state dependence in unemployment is strong and implies that an individual unemployed in the previous year is found to be more than twice as likely to be unemployed in the current year, compared to an individual employed in the previous year who has the same observed and unobserved characteristics. Finally, health limitations were found to increase the probability of unemployment in all approaches.

García-Gómez, Jones, and Rice (2010) used discrete-time hazard models to analyze the effects of deterioration of health on entries into and exits from employment,

³⁵ "Age4 = Age4/1,000,000)" in Hernández-Quevedo, Jones, and Rice (2005, p.14)

³⁶ The sample consisted on 3060 individuals, both women and men, who were classified either as employed or unemployed in 1991 and for whom information was available for the sample period 1991-1996.

using data from the first twelve waves of the British Household Panel Survey (1991–2002).³⁷ Physical health was measured as self-assessed health, general health conditions, self-reported health limitations, and two constructed latent indices of these measures in order to control for “reporting” bias. In addition, the GHQ index was included to measure psychological health.³⁸ García-Gómez et al. also used both initial period health and lagged health in the hazard models. The authors argued that by including initial period health and lagged health has led to alleviate the problem of endogeneity bias in the relationship between the health stock and employment transitions. Their empirical results indicated that health is a significant determinant of employment transitions, and that the effects for men are greater than those for women. Depending on the measure of health used, however, the hazard of becoming non-employed and the hazard of becoming employed vary. Poor physical poor health, calculated by health limitations and a constructed latent health index, has a substantial positive effect on exits from employment and a negative effect on entries into employment. Furthermore, poor psychological health tends to influence positively the hazard of becoming non-employed for the sample of individuals employed.

Brown et al. (2010) explored the effect of health on stated reservation wages of unemployed men and on market wages of employed men, using data from the first fourteen waves of the BHPS with particular interest in the potential endogeneity of health on labour market outcomes. They followed Stern (1989) and Bound (1991) and, along with the ordinal self-assessed health, they used information on more objective measures as

³⁷ The data consisted of individuals of working age who were classified as employed and non-employed. Employed are those who report to be either employed or self-employed. Non-employed, according to authors, are those who report unemployment, retirement, maternity leave, family care duties, being a student, being long-term sick or on a government training scheme. Women, in the sample, are aged between 16 and 59 years old.

³⁸ General health conditions were derived from questions related to specific health problems with: arms, legs or hands, sight, hearing, skin conditions or allergies, chest/breathing, heart/blood pressure, stomach or digestion, diabetes, anxiety or depression, alcohol or drugs, epilepsy or migraine, or other), the self-reported health limitations were derived from answer to the question “does your health in any way limit your daily activities compared to most people of your age?” and psychological health was measured upon questions related to concentration, sleep loss due to worry, perception of role, capability in decision making, whether constantly under strain, perception of problems in overcoming difficulties, enjoyment of day-to-of role, capability in decision making, whether constantly under strain, perception of problems in overcoming difficulties, enjoyment of day-to- day activities, ability to face problems, loss of confidence, self-worth, perception of problems in overcoming difficulties, enjoyment of day-to- day activities, ability to face problems, loss of confidence, self-worth, general happiness, and whether suffering depression or unhappiness.

instruments for the ordinal self-assessed health variable, along with socio-economic variables, to deal with the endogeneity issue. These measures are specific health problems, such as health problems with arms, legs, hands, sight, hearing, anxiety or depression and whether or not health limits daily activities. The self-assessed health was measured by responses to the following question: “In general would you say that your health is...very good, good, fair, poor, or very poor?” Brown et al. adopted the view that individuals with similar level of health may apply different thresholds when they report self-assessed health and that the same individual may even apply different thresholds over time. Therefore, they conducted their analysis with the use of generalized ordered probit approach, which allows for variations in self-assessed health thresholds, based on individuals’ characteristics. Their result indicated that poor health is a significant factor of whether or not the individual is attached to the labour market.

Recently, Jones, Rice, and Zantomio (2016) used data from the first five waves from Understanding Society, the successor of BHPS, to investigate how labour supply of individuals of working age responds to anticipated acute health shocks, measured by the incidence of cancer, stroke, and heart attack. Their empirical approach exploited innovations in health triggered by the advent of an acute health shock, occurring between the time periods $t - 1$ and t in order to identify the short-run labour supply response observed at time t , where acute health shock was treated as exogenous, conditional on observable characteristics and lagged outcomes. The objective health measure available in wave 1 and in subsequent waves used is a variable related to a question asking respondents whether they were diagnosed with specific health conditions, including cancer, heart attack, and stroke, in the past. This variable allowed the identification of individuals who have already experienced the onset of an acute health shock. In addition, more objective health measures were included in the analysis mostly relevant for cardiovascular disease, such as information about health risk factors, such as diabetes and high blood pressure and information about parents’ longevity (whether the mother and the father were alive when the respondent was aged 14). Finally, subjective health measures were present in the analysis including self-assessed health, the presence of a long-standing illness or disability and eleven types of limitations in daily activities. Their research showed clear evidence of considerable heterogeneity in observed responses to health shocks. Men are likely to reduce their labour supply by 6.4% after experiencing a health shock and women by 9.5%, with younger workers, as a whole, demonstrating stronger attachment to the labour market

after a health shock and, among older workers, those living with a partner demonstrating the weakest attachment compared to those who are single. This happens presumably, as the authors argued, because those living with a partner can rely on partner's financial support. For women in particular, older and more educated women significantly were found to reduce the number of hours worked after an acute health shock. In contrast, younger women not only did they not reduce hours worked, but they also reported increased labour market attachment (as measured by the desire to give up work or change employer). The authors ended up with the conclusion that financial constraints faced by less educated women, preferences, and intra-household division of labour and less opportunities to obtain alternative or less physically challenging jobs, have an important explanatory role to labour market exit for women (Jones et al., 2016).

More recently, Lenhart (2019), following a similar approach to Jones et al. (2016), examined the link between sudden health shocks and labour and household income, employment status and hours worked and, moreover, investigated whether increased health care expenditures and health care usage after a sudden health shock can explain the observed effects on such labour market outcomes. Into his investigation, Lenhart used data from waves 10–18 (2000–2008) of the BHPS to a sample of all individuals aged between 18 and 64. He measured sudden health shocks by self-assessed health, categorized into five categories from excellent to very poor, and by self-assessed responses to questions related to 15 health conditions (body pain, migraine, skin issues/allergy, asthma/chest pain, anxiety, heart or blood pressure, hearing problems, stomach/liver/kidney pain, seeing problems, epilepsy, diabetes, alcohol or drug problems, stroke, cancer or other conditions). His research showed evidence that deterioration of health leads to significant and persistent reductions in labour earnings for several years after the decline in health. The effects were found to be strongest for males, individuals with higher education and those working in managerial jobs. In the 5-year sample period, a health shock was shown to dramatically decrease labour earnings of men by £6,576.00, compared to a drop in labour earnings for women by £1,351.57. The difference effects between men and women, as Lenhart argues, could be partially explained by differences in pre-shock earnings (Lenhart, 2019).

Very recently, Cai (2021) explored the effect of health on the extensive and intensive margins of the labour supply in the UK with the use of the first seven waves of

the Understanding Society survey.³⁹ Both employment and hours worked exhibited positive state-dependence for both men and women. Cai (2021) employed a dynamic two-tiered tobit model and used the maximum simulated likelihood estimator (MSLE). In addition to self-assessed health, health conditions and disabilities, Cai used responses to questions related to the presence of new health conditions since the previous interview, such as asthma, arthritis and diabetes, hospitalization and length of stay in hospital and health lagged one year as an additional instrument⁴⁰ for health. All measures, except for lagged health, were assumed to be exogenous to labour supply. Cai (2021) showed evidence that health affects both the extensive and the intensive margins of labour supply for both genders, the effect of health is larger for the extensive margin of labour supply and that the effect of health for women dominates the effect of health for men. For women, specifically, a one unit increase in predicted health increases the probability of employment by 2.7 percentage points and hours worked conditional on being employed by 0.27 hour a week. The results also showed that when health is treated as exogenous, the estimated effects of health become smaller compared to the model in which health is treated as endogenous, leading to underestimation of the effect of health. The author argues that an explanation for the smaller effect could be attributed to measurement errors in self-assessed health. Similarly, when lagged health is excluded from the instruments, the estimated effects of health become much larger at both margins, which may lead to overestimating the effect of health on labour supply. Finally, education and the number of older children increases female labour supply at both the extensive and the intensive margins, whereas the number of children under school age and permanent non-labour income reduce labour supply at both margins.

³⁹ The sample consisted of individuals of working age between 21 and 64 years old, for whom information is available for at least two waves. In total, 12518 men and 16496 women constituted the final sample. For men, the employment rate ranged from 25.8 per cent for those reporting poor health to 89.6 per cent for those reporting excellent health and for women it ranged from 24.3 percent to 81.1 per cent respectively. Among employed men, the hours worked per week range from 36.5 hours for those with poor health to 39.3 hours for those with excellent health. For women the hours worked ranged from 28.4 to 30.5 hours per week respectively.

⁴⁰ Contoyannis et al. (2004) give evidence of strong positive state dependence of health which implies that past health should be a good predictor of current health.

Chapter 3

Theoretical Framework and Econometric Approaches

3.1 Introduction

Female labour supply can be considered as the outcome of a household's utility maximization, where a woman chooses her time allocation between work and leisure. In labour supply theory (see, for example Killingsworth, 1983), there is a critical wage, termed "reservation wage," which is considered to be the main determinant of whether the individual accepts a job offer or rejects it. In general, for wages below the reservation wage, the individual chooses to not participate in the labour force and receives, if eligible, a "reward" such as an unemployment benefit, whereas for wages above the reservation wage, the individual accepts the job offer. Thus, a woman's labour supply decision can be described as the difference between the utility she derives from accepting the job offer when working, and the utility from rejecting the job offer when not working. Our approximation for the female labour supply decision is in line with this concept and job search theory that emerged around 1970.⁴¹

In the next section we describe a dynamic utility maximization model of female labour supply that guides the empirical approaches that are described in section 3.3. Initially, we set up the basic components of the dynamic framework in Eckstein and

⁴¹ In her comprehensive review, Faggian (2014) argues that job search theory gained popularity in the 1970s as an alternative to the neoclassical "standard" labour supply theory. Faggian (2014) describes that the neoclassical "standard" framework didn't allow for unemployment, whereas individuals actively sought work, because it assumed perfect information. Consequently, under this framework, individuals were supposed to have only two choices, either being in employment or being out of the labour force. Subsequent empirical evidence suggested the significance of unemployment and, hence, an alternative theory was developed that could account for unemployment, known as "job search theory." The central principle of job search models, as Faggian (2014) states, is that job searching is a complex dynamic process, and it is up to individuals to determine when to stop this process under uncertainty and imperfect information (p. 60)

Wolpin (1989). However, since we estimate a reduced form specification⁴² of labour supply model rather than a structural one,⁴³ our approach more closely follows that of Hyslop (1999), who derives a dynamic binary labour force participation decision model that is consistent with life-cycle optimization. Specifically, based on Hyslop's (1999) dynamic utility maximization model of female labour supply, i.e., the decision to participate in the labor market emerges from the solution of a dynamic programming problem and depends on the difference between the reservation wage and the market wage, we construct a reduced form specification of labour supply. We substitute the reservation wage by a function of individual characteristics and market conditions, and we also substitute the market wage by a function of market conditions that determine the market wage. Hence, our approach is not entirely ad hoc, but is based on first principles, i.e., dynamic utility maximization under uncertainty. We also follow Michaud and Tatsiramos (2005) and Garibaldi and Wasmer (2003) who build their framework on Hyslop (1999).

3.2 A theoretical framework for intertemporal labour supply decision

Given her expectations, a married woman receives at each discrete period t a job offer and either accepts the job and participates in the labour market ($y_t = 1$) or rejects it and does not participate ($y_t = 0$). Her decision rule is determined by whether a latent variable, y_t^* , which reflects the difference in the expected payoffs of the $y_t = 1$ and $y_t = 0$ alternatives, crosses a threshold. The preferred alternative is the one with the largest difference in the expected payoff, which, without loss of generality, is zero. That is,

$$y_t = 1 \text{ if } y_t^* \geq 0,$$

⁴² A reduced-form equation is an equation that emerges from the solution of a system of behavioral equations and identities, which has an endogenous variable on the left hand side and exogenous or lagged endogenous variables on the right hand side.

⁴³ A structural equation is a behavioral equation or an identity describing reality. In our case, structural labour supply model refers to a model specification in which a woman's own wage, her partner's wage and probably working hours are included along with the couple's non-labour income.

$$y_t = 0 \text{ if } y_t^* < 0.$$

Due to the dynamic nature of our approach,⁴⁴ the latent variable y_t^* is the difference in alternative-specific value functions. The value function denotes the expected discounted value for each of two payoffs and is a function of previous labour supply decision, state variables, and unobservable factors.

Before we specify our empirical models, we follow closely Hyslop's (1999) theoretical analysis exposed in his section 2, in order to make clear that our empirical model is in "reduced form" (see Hyslop, 1999, p. 1262, and Keane and Sauer, 2009, p. 977).

According to Hyslop (1999), a married woman searches for a job in each period of her infinite lifetime. This search incurs a search cost. Assume she receives one job offer. Once she accepts it, this leads her to permanent employment at a fixed per-period wage w_t . Next period, she receives again a wage offer, but without searching. When a job offer is rejected, it cannot be recalled, and there is no on-the-job search. Hyslop (1999) finally assumes that work and home production (non-market work) are perfect substitutes. Therefore, there are only two alternatives for the labour supply decision, $y_t = 0$ and $y_t = 1$, and only two payoffs rewarded respectively.

Under the above assumptions and the assumption that she behaves rationally, a married woman maximizes her remaining discounted lifetime utility by choosing whether to participate ($y_{it} = 1$) or not to participate ($y_{it} = 0$). Hence, the objective of each woman is to maximize her expected present value of discounted utility over an infinite lifetime

$$(3.1) \quad U_t = \sum_{s=0}^{\infty} \frac{1}{(1 + \rho)^s} E_t u(C_{t+s}, y_{t+s}, \Omega_{t+s})$$

with respect to the dichotomous variable y_t , where $s = 0$ is the theoretical start of the decision process in an infinite time horizon, $u(\cdot)$ is household's period flow utility at time

⁴⁴ The dynamics of the decision process emerge from the dependence of current decisions on previous decisions.

t , C_t stands for household's consumption during period t , Ω_t is the vector of characteristics of the family in period t , and ρ is the rate of time preference. The expectations operator E_t in (3.1) summarizes a married woman's expectations about future values of the variables, i.e. income, value of home production, and taste shifters.

The period flow utility function for a married woman i at time t , is given by

$$(3.2) \quad u = u(C_t, y_t, \Omega_t).$$

The period flow utility function is defined over the joint family consumption C_t , the woman's own participation state y_t , and the vector of characteristics of the family Ω_t in period t . Husband's participation state is assumed exogenous.

The household's budget constraint in each calendar year t is specified as

$$(3.3) \quad C_t = m_t + w_t y_t - \gamma_1(1 - y_{t-1}),$$

where m_t is non-labour income, w_t is the wage, y_t and y_{t-1} are the participation states in periods t and $t - 1$, respectively, and γ_1 is the cost of search, which measures information search costs and opportunity cost of time devoted to search. Hyslop, by assumption, considers that both cost of search and the wage distribution are time-invariant and the same for all women. Consequently, Hyslop ignores the choice of hours of work and treats labour supply as a discrete choice. Note, also, that Hyslop (1999), as well as Michaud and Tatsiramos (2005), ignore accumulation of wealth (human and nonhuman) and borrowing/lending and assume a static budget constraint. That is, the household consumes all of its income within time period t .

We follow the standard methodology of dynamic programming. The value function at the beginning of period t is defined as

$$(3.4) \quad V(y_{t-1}, \Omega_t) = \max(V^0(y_{t-1}, \Omega_t), V^1(y_{t-1}, \Omega_t)),$$

where V^1 and V^0 represent maximum expected discounted utility in period t if the woman participates in period t (state 1) or does not participate in period t (state 0), respectively. For $y_{t-1} = 0$, that is, for a nonparticipant woman at $t - 1$ period, the value function at time t is defined as

$$(3.5) \quad V(0, \Omega_t) = \max(V^0(0, \Omega_t), V^1(0, \Omega_t)).$$

Comparing $V^0(0, \Omega_t)$ with $V^1(0, \Omega_t)$ implies that there exists a reservation wage w_{0t}^* for which a nonparticipant woman is indifferent between working and not working. This reservation wage is defined by

$$(3.6) \quad V^0(0, \Omega_t) = V^1(0, \Omega_t | w_{0t}^*),$$

or

$$(3.7)$$

$$u(y_t + w_{0t}^* - \gamma_1, 1, \Omega_t) + \frac{1}{1+\rho} E_t V(1, \Omega_{t+1}) = u(y_t - \gamma_1, 0, \Omega_t) + \frac{1}{1+\rho} E_t V(0, \Omega_{t+1}).$$

The left-hand side of (3.7) is Bellman's equation of a nonparticipant in period $t - 1$ for state 1 (participation) in period t , whereas its right-hand side is Bellman's equation of a nonparticipant in period $t - 1$ for state 0 (non-participation) in period t . The woman participates at time t if $V^1(0, \Omega_t | w_{0t}^*)$ is greater than $V^0(0, \Omega_t)$ or, equally, if the market wage w_t exceeds the reservation wage w_{0t}^* , that is, if $w_t > w_{0t}^*$. For $y_{t-1} = 1$, that is, for a participant woman at time $t - 1$, the value function at time t is defined as

$$(3.8) \quad V(1, \Omega_t) = \max(V^0(1, \Omega_t), V^1(1, \Omega_t)).$$

Comparing $V^0(1, \Omega_t)$ with $V^1(1, \Omega_t)$ implies that a participant at time $t - 1$ is indifferent between participating and not participating at time t given her reservation wage w_{1t}^* , such that

$$(3.9) \quad V^0(1, \Omega_t) = V^1(1, \Omega_t | w_{1t}^*),$$

or

$$(3.10) \quad u(y_t + w_{1t}^*, 1, \Omega_t) + \frac{1}{1+\rho} E_t V(1, \Omega_{t+1})$$

$$= u(y_t, 0, \Omega_t) + \frac{1}{1 + \rho} E_t V(0, \Omega_{t+1}).$$

Equation (3.10) is similar to (3.7). Its left-hand side is Bellman's equation of a participant in period $t - 1$ for state 1 (participation) in period t , whereas its right-hand side is Bellman's equation of a participant in period $t - 1$ for state 0 (non-participation) in period t . A participant woman at time $t - 1$ will also participate at time t if $V^1(1, \Omega_t | w_{1t}^*)$ is greater than $V^0(1, \Omega_t)$ or, equally, if the market wage w_t exceeds the reservation wage w_{1t}^* , that is, if $w_t > w_{1t}^*$.

Subtracting (3.7) from (3.10) yields

$$(3.11) \quad \begin{aligned} & u(y_t + w_{1t}^*, 1, \Omega_t) - u(y_t + w_{0t}^* - \gamma_1, 1, \Omega_t) \\ &= u(y_t, 0, \Omega_t) - u(y_t - \gamma_1, 0, \Omega_t). \end{aligned}$$

Taylor series expansions (see Appendix [A.1.1](#) for a proof) of the left and right hand sides around $y_t + w_{0t}^*$ and y_t , respectively, give the following connection between the two reservation wages as

$$(3.12) \quad w_{1t}^* \approx w_{0t}^* - \gamma_1 \left(1 - \frac{u_1(y_t, 0, \Omega_t)}{u_1(y_t + w_{0t}^*, 1, \Omega_t)} \right) = w_{0t}^* - \gamma,$$

where

$$\gamma = \gamma_1 \left(1 - \frac{u_1(y_t, 0, \Omega_t)}{u_1(y_t + w_{0t}^*, 1, \Omega_t)} \right),$$

and $u_1(\cdot)$ is the partial derivative of the utility function with respect to consumption. The numerator of the fraction in (3.12) indicates the marginal utility of consumption when not

participating and the denominator indicates the marginal utility of consumption when participating. Note also that $w_{1t}^* = w_{0t}^* = w_t$ when $\gamma = 0$.⁴⁵

Hence, conditional on the reservation wage, the family characteristics, and preference shocks, the decision rule for participation in period t can be described by

$$(3.13) \quad y_t = 1(w_t > w_{0t}^* - \gamma y_{t-1}),$$

where $1(\cdot)$ is an indicator function that equals 1 if the expression inside the parenthesis is true and 0 otherwise. The parameter γ is a key parameter in this thesis, as it captures true state-dependence (Heckman, 1981a).

In the literature that explores female dynamic labour supply, the estimation of a reduced-form specification is justified for two main reasons. First, we do not observe wages for women who do not work. Second, even if we observed wages, endogeneity could arise due to a possible correlation between wages and other explanatory variables or unobserved factors. Consequently, wages are not included in a reduced-form equation as an explanatory variable (see, for example, Tatsiramos, 2008).

In the next section, we present our econometric methodology and examine empirical specifications of female labour supply, which accounts for linkages between labour supply decisions and key demographic characteristics, with emphasis on health. Our estimation will be carried out with household panel data from the UK.

⁴⁵ Garibaldi and Wasmer (2003) demonstrate the existence of the inequality $w_{1t}^* < w_t < w_{0t}^*$ where $w_t = w(1 - n)$ is the gross wage w net of taxes n . In their work, w_{1t}^* is lower than w_t because employed individuals face uncertainty about future incomes and under this uncertainty they prefer staying in work rather than being unemployed, even if it is possible that home production produces more consumption than market work. This assumption justifies a lower reservation wage w_{1t}^* than the market wage w_t .

3.3 Econometric approaches

In this section, following Lucchetti and Pignini (2017), we detail our empirical econometric strategy for estimating a reduced-form equation for labour supply described in the previous section. We describe the connection between the reduced-form specification and binary dynamic probit models. We analytically present alternative approaches for the probability of female labour force participation proposed by Heckman (1981b), Hyslop (1999), Keane and Sauer (2009) and Bartolucci and Nigro (2010). We also employ Mundlak's (1978) approach, embedded in Hyslop's (1999) approach, to test for potential endogeneity of health-related variables.

The equation for the latent dependent variable is specified as

$$(3.14) \quad y_{it}^* = \gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i + \varepsilon_{it}, \quad i = 1, \dots, n, \quad t = 2, \dots, T$$

where y_{it}^* is the latent dependent variable, which is defined in Section 3.2 and which is unobservable. It may be interpreted as utility or propensity of the labour supply decision. The dummy variable y_{it} , which is also defined in Section 3.2, is the observed binary outcome for the participation state that is assumed to maximize utility. The subscript i denotes individuals and the subscript t denotes time periods.

In equation (3.14), $\mathbf{x}_{it}$ denotes a vector of explanatory variables whose first element is 1, $\boldsymbol{\beta}$ is the vector of the regression parameters associated with the explanatory variables, a_i is the time invariant unobserved individual heterogeneity that accounts for the time invariant individual effects, and ε_{it} is the time-varying unobserved error term. The error term, also called idiosyncratic error, is assumed independent of the explanatory variables $\mathbf{x}_{it}$ and the individual heterogeneity a_i . We assume $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$. The unobserved individual heterogeneity a_i can be assumed to be either a random variable that is uncorrelated with all explanatory variables $\mathbf{x}_{it}$, or a fixed parameter to be estimated.

The parameters of main interest in equation (3.14) are typically $\boldsymbol{\beta}$ for the explanatory variables and γ for the true state dependence. These are referred to as

structural parameters. The time invariant individual effects a_i are referred to as incidental parameters and ignoring them may lead to inconsistent estimates and upwards biased state dependence effect (Bartolucci and Pignini, 2017; Stewart, 2006).

If we assume that the estimate of state dependence parameter γ is different from zero and that the unobserved individual heterogeneity a_i is present, direct estimation of the model (3.14) suffers from the initial conditions problem; we lack knowledge of the data generating process that governs the initial response. Consequently, there might be correlation between the lagged dependent variable $y_{i,t-1}$ and the unobserved individual heterogeneity, yielding inconsistent estimates (Michaud and Tatsiramos, 2008, p. 11) and an overstated estimate of state dependence (Stewart, 2006, p. 258). Therefore, estimation of equation (3.14) requires an assumption about the initial observations y_{i1} and, specifically, how these observations are related to the individual heterogeneity. Specifically, for our sample, this “initial conditions problem” occurs because our analysis starts at the age of 39 for women and these observations for participation state do not necessarily coincide with the start of the participation status process, say at age 18. Indeed, we could not argue that a woman’s participation status at the age of 39 is unrelated to her individual time invariant effects. Consequently, we could not assume that the initial observations y_{i1} are exogenous.

Therefore, the process for the response variable must be initialized to explain how participation state relates to the process before the observations are available. As we have described in the review of the literature (Chapter 2), the dependency of participation state on the past is introduced by the response variable lagged one period as a regressor in female labour supply models.

3.3.1 Random-Effects Approaches

The following subsections 3.3.1.1, 3.3.1.2 and 3.3.1.3 are in line with Lucchetti and Pignini (2017). The justification for doing so is that we use the software developed by Lucchetti and Pignini (Gretl), so the models should be in suitable form.

3.3.1.1 The Lucchetti and Pignini's (2017) version of Heckman's (1981) estimator

Heckman (1981b) was the first to suggest a solution strategy to approximate the initial conditions problem. Heckman (1981b) proposed a solution to the initial conditions problem, which involves a simultaneous estimation of two equations, a “structural equation” and a “reduced-form equation.” For the years following the initial year, the structural equation is used to estimate the probability of participating in the current year as a function of the participation status in the previous year. The reduced form equation accounts for the initial conditions and is used to estimate the probability of participating in the first year of the sample.

Thus, Heckman's (1981b) estimator follows the typical formulation of a dynamic random effects discrete choice model with an additional equation for the initial observation y_{i1} (Lucchetti and Pignini, 2017, p.5).

$$(3.15) \quad y_{it} = \mathbf{1} \{ \gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i + \varepsilon_{it} \geq 0 \} \quad \text{for } i = 1, \dots, n, \quad t = 2, \dots, T$$

$$(3.16) \quad y_{i1} = \mathbf{1} \{ \mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i + \varepsilon_{i1} \geq 0 \} \quad \text{for } i = 1, \dots, n$$

where y_{it} is the binary response variable, $\mathbf{1}\{\cdot\}$ is an indicator function, $\mathbf{x}_{it}$ is a vector of explanatory variables and $\mathbf{z}_{i1}$ contains the explanatory variables only in the first period. Equation (3.16) informs us about the estimation of equation (3.15) because of the presence of unobserved individual heterogeneity a_i . Further, equation (3.16) can be assumed as the linear approximation of utility of chosen participation state at time $t = 1$.

Exogeneity of the initial conditions can be considered as resulting from imposing the restriction $\theta = 0$ on equation (3.16). The rejection would imply that the time-invariant unobserved individual heterogeneity α_i is correlated with the initial period. The assumptions on α_i and ε_i are:

- (i) $E[\varepsilon_{it}|\mathbf{X}_i, \alpha_i] = 0$ (independence between α_i and ε_i);
- (ii) $E[\alpha_i|\mathbf{X}_i] = 0$ (orthogonality between α_i and $\mathbf{X}_i$);
- (iii) $E[\varepsilon_{it}\varepsilon_{is}] = 0$ for $t, s = 1, \dots, T$, $t \neq s$ (absence of autocorrelation in ε_i); and
- (iv) $[\theta\alpha_i + \varepsilon_{i1}, \alpha_i + \varepsilon_{i2}, \dots, \alpha_i + \varepsilon_{iT}]^T \sim N(\mathbf{0}; \boldsymbol{\Sigma})$ (joint normality conditional on $\mathbf{X}_i$),

where $V(\alpha_i) = \sigma_\alpha^2$, $V(\varepsilon_{it}) = 1$ for all t (normalization, as in Heckman, 1981b, p. 181), and

$$(3.17) \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1 + \theta^2 \sigma_\alpha^2 & \theta \sigma_\alpha^2 & \theta \sigma_\alpha^2 & \dots \\ \theta \sigma_\alpha^2 & 1 + \sigma_\alpha^2 & \sigma_\alpha^2 & \dots \\ \theta \sigma_\alpha^2 & \sigma_\alpha^2 & 1 + \sigma_\alpha^2 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix},$$

as $V(\theta\alpha_i + \varepsilon_{i1}) = 1 + \theta^2 \sigma_\alpha^2$, $E(\theta\alpha_i + \varepsilon_{i1})(\alpha_i + \varepsilon_{i2}) = \theta \sigma_\alpha^2$, etc.

Under these assumptions, the parameter vector $\psi = [\boldsymbol{\beta}^T, \gamma, \boldsymbol{\pi}^T, \theta, \sigma_\alpha]$ can be estimated by maximum likelihood.

The following expression gives the i th contribution to the likelihood (see Appendix [A.1.2](#) for a proof)

$$\mathcal{L}_i(\psi) = \int_{-\infty}^{\infty} \Phi[(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta \alpha_i)(2y_{i1} - 1)] \prod_{t=2}^T \Phi[(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \alpha_i)(2y_{it} - 1)] d\Phi\left(\frac{\alpha_i}{\sigma_\alpha}\right),$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Under the assumption that the unobserved heterogeneity α_i is normally distributed, the above integral is taken

over α_i . Heckman's approach can be evaluated by means of Gauss-Hermite quadrature and estimation is carried out by maximum likelihood.

Hence, Heckman (1981b) solves the initial conditions problem, accounts for unobserved individual heterogeneity, and distinguishes "true" state dependence from "spurious" state dependence.

3.3.1.2 Lucchetti and Pignini's (2017) version of Hyslop's (1999) estimator

Hyslop (1999) generalized Heckman's approach by allowing for autocorrelation in ε_{it} . Heckman (1981b) isolated the true state dependence, captured by γ in equation (3.15), from persistence induced by time-invariant unobserved individual heterogeneity α_i in equation (3.15). Hyslop, further, attempted to isolate the true state dependence from both individual heterogeneity α_i and serial correlation in ε_{it} . That is, the persistence in the time-varying unobserved errors is also parameterized.

Serial correlation could be the outcome of omitted variables that are autocorrelated, but is usually taken to mean that there are transitory shocks with effects that last more than one period. If there is serial correlation that is not accounted for, the coefficient of the lagged labour supply variable is biased downwards, thus the state dependence is underestimated (see, for example, Tatsiramos, 2011, p.17).

Hyslop (1999) assumes that the error term ε_{it} follows an AR(1) process

$$(3.18) \quad \varepsilon_{it} = \rho\varepsilon_{i,t-1} + \eta_{it} \text{ for } t = 2, \dots, T$$

where $|\rho| < 1$ and $\eta_{it} \sim N(0, 1 - \rho^2)$. Therefore, the variance-covariance matrix of the error components is modified as follows:

$$(3.19) \quad \Sigma = \begin{bmatrix} 1 + \theta^2 \sigma_\alpha^2 & \rho + \theta \sigma_\alpha^2 & \rho^2 + \theta \sigma_\alpha^2 & \dots \\ \rho + \theta \sigma_\alpha^2 & 1 + \sigma_\alpha^2 & \rho + \sigma_\alpha^2 & \dots \\ \rho^2 + \theta \sigma_\alpha^2 & \rho + \sigma_\alpha^2 & 1 + \sigma_\alpha^2 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

Notice that for $\rho = 0$, equation (3.19) reduces to equation (3.17).

3.3.1.3 The Lucchetti and Pignini's (2017) version of Keane and Sauer's (2009) estimator

Keane and Sauer (2009) further generalized Hyslop (1999) and introduced a more flexible treatment of the initial conditions equation (i.e., the “reduced form equation”) in which they defined an additional parameter, τ .

Hence, equation (3.16) becomes

$$(3.20) \quad y_{i1} = \mathbf{1} \{ \mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i + u_i \geq 0 \}$$

with $E(u_i \cdot \varepsilon_{i2}) = \tau$. Lucchetti and Pignini (2017) argue that the additional parameter, τ , is effectively a correlation coefficient. Note also that both u_i and ε_{it} must be normalized for identification (Lucchetti and Pignini, 2017, p. 6). Therefore, equation (3.19) becomes $[\theta a_i + u_i, \alpha_i + \varepsilon_{i2}, \dots, \alpha_i + \varepsilon_{iT}]^T \sim N(\mathbf{0}; \Sigma)$ with

$$(3.21) \quad \Sigma = \begin{bmatrix} 1 + \theta^2 \sigma_\alpha^2 & \tau \rho + \theta \sigma_\alpha^2 & \tau \rho^2 + \theta \sigma_\alpha^2 & \dots \\ \tau \rho + \theta \sigma_\alpha^2 & 1 + \sigma_\alpha^2 & \rho + \sigma_\alpha^2 & \dots \\ \tau \rho^2 + \theta \sigma_\alpha^2 & \rho + \sigma_\alpha^2 & 1 + \sigma_\alpha^2 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

For $\tau = 1$, equation (3.21) reduces to equation (3.19) and for $\rho = 0$ equation (3.21) reduces to equation (3.17).

Estimation of the binary probit models of Hyslop (1999) and Keane and Sauer (2009), which account for individual heterogeneity and serial correlation among idiosyncratic errors, requires multiple integration. The simulated maximum likelihood method can surpass the computation burden in the estimation process of choice probabilities. By using the GHK algorithm,⁴⁶ we compute the likelihood function:

$$(3.22) \quad \mathcal{L}_i^*(\psi) = \frac{1}{R} \sum_{r=1}^R \Phi_{Tr}^*(\alpha_i, \mathbf{b}_i, \mathbf{C})$$

with $\alpha_i = (\mathbf{z}_{i1}^T \boldsymbol{\pi})(2y_{i1} - 1)$ and $\mathbf{b}_i = [b_{i2}, \dots, b_{iT}]$, where $b_{it} = (\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta})(2y_{it} - 1)$, $\mathbf{C}$ is the lower-triangular Cholesky factor of $\boldsymbol{\Sigma}$ (i.e., $\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}'$), and R is the number of random draws used in the simulation (see Keane and Sauer, 2006, pp. 7-8; Stewart, 2007, pp.517-578).

In our analysis in chapter 5, we also use simulated maximum likelihood for Heckman's approach in order to be in line with simulated maximum likelihood estimation performed for Hyslop (1999) and Keane and Sauer (2009) approaches.

3.3.2 Fixed-Effects Approach

Apart from the random-effects approach, another way to surmount the initial conditions problem is the fixed-effects approach,⁴⁷ which permits the consistent estimation of the parameters with no distributional assumptions on the unobserved individual heterogeneity a_i .

⁴⁶ The GHK simulator has been developed after works by Geweke (1988), Hajivassiliou (1993), and Keane (1994).

⁴⁷ The presentation of the fixed-effects approach is based on Bartolucci and Pignini (2017) and on Lucchetti and Pignini (2017).

The key idea is to condition the joint distribution of $\mathbf{y}_i$ on a suitably defined sufficient statistic for α_i , where $\mathbf{y}_i$ is the overall vector of response variables, defined as $\mathbf{y}_i = [y_{i1}, \dots, y_{iT}]$. For the static fixed-effects logit model, that is the fixed-effects logit model without the lagged response variable $y_{i,t-1}$ among the explanatory variables, it is feasible to discard the individual heterogeneity α_i by conditioning on simple sufficient statistics.⁴⁸ The estimator based on this approach is commonly known as Conditional Maximum Likelihood (CML) estimator. The inclusion of the lagged response variable $y_{i,t-1}$ among the explanatory variables, however, does not allow a simple sufficient statistic for the individual heterogeneity α_i and, therefore, the model cannot be estimated by CML. This drawback is surmounted by Bartolucci and Nigro (2010), who developed a model based on a Quadratic Exponential (QE) formulation set out in Cox (1972) for analyzing dynamic discrete choice panel data models. Their approach results in eliminating the individual heterogeneity α_i by conditioning the likelihood $p(\mathbf{y}_i | \mathbf{X}_i, \alpha_i; \psi)$ on sufficient statistics y_{i+} for the parameters of unobserved heterogeneity, which correspond to the sums of the response variables at the individual level. The parameter vector ψ is defined as $\psi = [\boldsymbol{\beta}^T, \gamma, \boldsymbol{\pi}^T, \theta, \sigma_\alpha]$. As a result, the initial condition does not need to be dealt with, so that the joint probability can be written as $p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, y_{i+}; \psi)$. Therefore, the parameters of the model can be easily estimated by the CML method and are $\sqrt{n}$ consistent.

However, in empirical research, this approach has not been generally adopted because it requires that at least a transition between the states 0 and 1 is observed for the individual to contribute to the likelihood.⁴⁹ As a result, the number of usable observations often decreases drastically compared to the sample size, especially in cases where strong persistence in the dependent variable is precisely the reason why a dynamic model is needed. Nevertheless, we also employ this approach because it gives a reliable inference in case where we suspect that there may be correlation between the individual

⁴⁸ When the lagged response variable is not included in the model and thus true state dependency is not taken into account, the sums of the response variables at the individual level are sufficient statistics for the incidental parameters. These sufficient statistics are referred to as total scores (Bartolucci and Pigini, 2017, p. 3).

⁴⁹ See the example of Baltagi (2005, chapter 11, page 210) that demonstrates how observations, where the dependent variable does not change over time, add nothing to the conditional log likelihood.

heterogeneity a_i and at least one explanatory variable (Lucchetti and Pignini, 2017, top of p.3).

We first describe the CML method applied to the static logit model because it is the basic framework, then we present the dynamic logit and finally the QE formulation by Bartolucci and Nigro (2010).

Let y_{it} be a sequence of binary responses referred to the same individual i at time t with $i = 1, \dots, n$ and $t = 1, \dots, T$. For the static logit model, the conditional distribution of a single response is (see Appendix [A.1.3](#) for a derivation)

$$(3.23) \quad p(y_{it}|X_i, \alpha_i; \psi) = \frac{\exp[y_{it}(x_{it}^T \beta + a_i)]}{1 + \exp(x_{it}^T \beta + a_i)},$$

where a_i is the individual heterogeneity and β is the vector of regression parameters associated with the explanatory variables x_{it} .

For the joint probability of $\mathbf{y}_i = (y_{i1}, \dots, y_{iT})^T$, the static logit model yields the following likelihood function (see Appendix [A.1.4](#) for a derivation):

$$(3.24) \quad p(\mathbf{y}_i|X_i, \alpha_i; \psi) = \frac{\exp(\sum_t y_{it} x_{it}^T \beta) \exp(a_i y_{i+})}{\prod_t [1 + \exp(x_{it}^T \beta + a_i)]},$$

where the summation and product operators, $\sum_t$ and $\prod_t$, range over $t = 1, \dots, T$, and $y_{i+} \equiv \sum_t y_{it}$ is the total score. Andersen (1970), as cited in Bartolucci and Pignini (2017, p. 3) shows that y_{i+} is a sufficient statistic for a_i , i.e., the joint probability of $\mathbf{y}_i$, conditional on y_{i+} , does not depend on a_i . That is (see Appendix [A.1.5](#) for the proof),

$$(3.25) \quad p(\mathbf{y}_i|X_i, y_{i+}; \psi) = \frac{p(\mathbf{y}_i|X_i, \alpha_i; \psi)}{p(y_{i+}|X_i, \alpha_i; \psi)},$$

where the denominator contains only the vectors of binary responses $\mathbf{b} = (b_1, \dots, b_T)^T$ such that $b_+ = y_{i+}$, where $b_+ = \sum_t b_t$. Hence, equation (3.25) becomes

$$(3.26) \quad p(\mathbf{y}_i | \mathbf{X}_i, y_{i+}; \boldsymbol{\psi}) = \frac{p(\mathbf{y}_i | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})}{\sum_{\mathbf{b}: b_+ = y_{i+}} p(\mathbf{b} | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})},$$

with

$$(3.27) \quad p(\mathbf{b} | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi}) = \frac{\exp(\sum_t b_t \mathbf{x}_{it}^T \boldsymbol{\beta}) \exp(a_i b_+)}{\pi_t [1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}.$$

Therefore, substituting Equations (3.24) and (3.27) into Equation (3.26), the conditional distribution of the vector of responses $\mathbf{y}_i$ is

$$(3.28) \quad \begin{aligned} & p(\mathbf{y}_i | \mathbf{X}_i, y_{i+}, \alpha_i; \boldsymbol{\psi}) \\ &= \frac{\exp(\sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta}) \exp(a_i y_{i+})}{\pi_t [1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]} \frac{\pi_t [1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}{\sum_{\mathbf{b}: b_+ = y_{i+}} \exp(a_i b_+) \exp(\sum_t b_t \mathbf{x}_{it}^T \boldsymbol{\beta})} \\ &= \frac{\exp(\sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta})}{\sum_{\mathbf{b}: b_+ = y_{i+}} \exp(\sum_t b_t \mathbf{x}_{it}^T \boldsymbol{\beta})} = p(\mathbf{y}_i | \mathbf{X}_i, y_{i+}; \boldsymbol{\psi}), \end{aligned}$$

where the individual heterogeneity α_i has been canceled out. The conditional log-likelihood based on the above distribution can be written as

$$(3.29) \quad l(\boldsymbol{\psi}) = \sum_{i=1}^n I\{0 < y_{i+} < T\} \log p(\mathbf{y}_i | \mathbf{X}_i, y_{i+}; \boldsymbol{\psi}),$$

where the indicator function $I\{\cdot\}$ is implemented to take into account that observations whose total score is 0 or T do not contribute to the likelihood. Then, a Newton-Raphson

algorithm maximizes this conditional log-likelihood with respect to ψ , obtaining the $\widehat{\psi}$ CML estimator.

For the dynamic logit model, the conditional density of y_{it} given $p(y_{it} | y_{i,t-1}, \mathbf{X}_i, \alpha_i; \psi)$, is (see Appendix [A.1.6](#) for a proof)

$$(3.34) \quad p(y_{it} | y_{i,t-1}, \mathbf{X}_i, \alpha_i; \psi) = \frac{\exp[y_{it}(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}{1 + \exp(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)}.$$

Note that equation (3.34) is an extension of equation (3.23), which is augmented to include $y_{i,t-1}$ amongst the explanatory variables. Here, the individual heterogeneity a_i is viewed as a fixed parameter, and the initial observation can be treated as given.

Hence, the joint probability of the overall vector of response variables $\mathbf{y}_i$, conditional on the initial observation y_{i1} , is (see Appendix [A.1.7](#) for a proof)

$$(3.31) \quad p(\mathbf{y}_i | y_{i1}, \mathbf{X}_i, \alpha_i; \psi) = \frac{\exp(y_{i*} \gamma + \sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta} + y_{i1} a_i)}{\prod_t [1 + \exp(y_{i,t-1} \gamma + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]},$$

where $y_* = \sum_t y_{i,t-1} y_{it}$ and sums and products go from $t = 2$ to T .

The dynamic logit model, unlike the static logit model of equation (3.23), does not admit sufficient statistics for individual heterogeneity a_i . As a result, CML inference is not easily implementable; it can only be derived in the case of $T = 3$ and when there are no explanatory variables in the model (Bartolucci and Pignini, 2017, p.5). The shortcomings of the fixed-effects dynamic logit model are surmounted by the approximating QE model developed in Bartolucci and Nigro (2010), which defines the conditional distribution of $\mathbf{y}_i$ as follows (see Bartolucci and Pignini, 2017, p. 5):

(3.32)

$$p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi}) = \frac{\exp[y_{i+}\gamma + \sum_t y_{it} x_{it}^T \boldsymbol{\beta}_1 + y_{iT}(\mu + x_{iT}^T \boldsymbol{\beta}_2) + y_{i+}\alpha_i]}{\sum_{\mathbf{b} \in B} \exp[\sum_t b_t b_{t-1} \gamma + \sum_t b_t x_{it}^T \boldsymbol{\beta}_1 + b_T(\mu + x_{iT}^T \boldsymbol{\beta}_2) + b_+ \alpha_i]}$$

where $b_+ \equiv \sum_t b_t$ and $B \equiv \{\mathbf{b}: \mathbf{b} \in \{0,1\}^T\}$, that is the vectors of binary responses $\mathbf{b}$ that are equal to zero or one (Lucchetti and Pignini, p.7).

Unlike the dynamic logit model, the QE model admits sufficient statistics for the individual heterogeneity α_i , which are removed by conditioning on the total score y_{i+} . In particular, following the same derivations as for the conditional maximum likelihood estimation for the standard fixed-effects logit model, Equations (3.23)-(3.28), we obtain (see Appendix [A.1.8](#) for a proof)

(3.33)

$$p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, y_{i+}; \boldsymbol{\psi}) = \frac{p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi})}{p(y_{i+} | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi})} = \frac{\exp[\sum_t y_{it} y_{i,t-1} \gamma + \sum_t y_{it} x_{it}^T \boldsymbol{\beta}_1 + y_{iT}(\mu + x_{iT}^T \boldsymbol{\beta}_2)]}{\sum_{\mathbf{b}: b_+ = y_{i+}} \exp[\sum_t b_t b_{t-1} \gamma + \sum_t b_t x_{it}^T \boldsymbol{\beta}_1 + b_T(\mu + x_{iT}^T \boldsymbol{\beta}_2)]}$$

Notice that the right-hand-side of (3.33) does not depend on α_i ; the parameters for the unobserved heterogeneity have been removed by conditioning on the total score, y_{i+} . Hence, y_{i+} is a sufficient statistic for the unobserved heterogeneity in equation (3.33) which contains only those vectors $\mathbf{b} \in B$ such that $b_+ = y_{i+}$. Consequently, the conditional log-likelihood can be expressed as

$$(3.34) \quad l(\boldsymbol{\psi}) = \sum_{i=1}^n 1\{0 < y_{i+} < T\} \log p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, y_{i+}; \boldsymbol{\psi})$$

and is maximized with respect to the parameter vector $\boldsymbol{\psi} = [\gamma, \boldsymbol{\beta}_1^T, \mu, \boldsymbol{\beta}_2^T]^T$. The maximization of the conditional log-likelihood may be performed by a Newton- Raphson

algorithm and the parameter vector ψ is $\sqrt{n}$ consistent and has asymptotic normal distribution (Bartolucci and Pignini, 2017, p. 6).

Moreover, Bartolucci and Nigro (2010, p. 723) show that for $t = 2, \dots, T$ the parameter γ has the same interpretation that it has under the dynamic logit model, that is, log-odds ratio between each pair of consecutive response variables y_{it} s. If we substitute for $y_{it} = 1$ and $y_{it} = 0$ from Bartolucci and Nigro (2010, p.721, equation (2)), we have

(3.35)

$$\log \frac{p(y_{it} = 1|\alpha_i, \mathbf{X}_i, y_{i,t-1} = 1)p(y_{it} = 0|\alpha_i, \mathbf{X}_i, y_{i,t-1} = 0)}{p(y_{it} = 0|\alpha_i, \mathbf{X}_i, y_{i,t-1} = 1)p(y_{it} = 1|\alpha_i, \mathbf{X}_i, y_{i,t-1} = 0)} = \gamma.$$

In the fixed effect framework when the assumed logit model includes the lagged response variable, another conditional maximum likelihood estimator is proposed by Honoré and Kyriazidou (2000), on the basis of a weighted conditional log-likelihood. However, their estimation process imposes strong requirements on the distribution of the covariates. In particular, it cannot include discrete covariates, such as age and time dummies, it demands a non-negligible computational burden and, although the estimator is consistent, its rate of convergence to the true parameter value is slower than $\sqrt{n}$ (Bartolucci and Pignini, 2017, p. 5).

3.3.3 Mundlak's Approach

All the random-effects models in section 3.3.1 assume that the unobserved individual heterogeneity a_i is uncorrelated with all the explanatory variables $\mathbf{x}_{it}$. To test for endogeneity in a non-linear framework, Mundlak (1978) specified an approach that allows the unobserved individual heterogeneity a_i to be correlated with the explanatory variables and his correlated random effects approach is considered to be the parametric counterpart of the fixed-effects approach (see Wooldridge, 2010, Chapter 15.8.2, p.615).

Mundlak (1978) specified a relationship between the unobserved individual heterogeneity a_i and all the explanatory variables through a linear form as

$$(3.36) \quad a_i = \bar{\mathbf{x}}_i \mathbf{c} + \xi_i,$$

where ξ_i is i.i.d. $\sim N(0, \sigma_\xi^2)$ and independent of $\mathbf{x}_{it}$ and ε_{it} , for all i, t and where ε_{it} is i.i.d. $N(0, 1)$; see equation (3.15). The vector $\bar{\mathbf{x}}_i$ includes the time averages of all time-varying explanatory variables. This approach suggests that equation (3.15) be augmented by additional time invariant regressors $\bar{\mathbf{x}}_i$, i.e.,

$$(3.37) \quad y_{it} = 1 \{ \gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i + \varepsilon_{it} \geq 0 \}, \quad t = 2, \dots, T.$$

Hence, the unobserved individual heterogeneity is uncorrelated with the explanatory variables if and only if $\mathbf{c} = \mathbf{0}$ (see Baltagi, 2005, Chapter 7, p.125).

This specification implies that the intra-group correlation coefficient (otherwise cross-period correlation for the composite error term) between $\rho_{it} = \xi_i + \varepsilon_{it}$ in any two different periods will be the same: $\rho = \text{corr}(\rho_{it}, \rho_{is}) = \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\varepsilon^2}$ for $t, s = 2, \dots, T, t \neq s$. Notice that under a normalization requirement, since y is a binary variable, we consider $\sigma_\varepsilon^2 = 1$ (see, for example, Stewart, 2007, p. 515). Therefore, our model now controls for unobserved individual heterogeneity ξ_i through correlated random effects and should therefore remove potential bias from the estimate of state dependence. As in the fixed-effects approach by Bartolucci and Nigro (2010) in section 3.3.2, the time-invariant explanatory variables are excluded from the model specification, in order to avoid perfect correlation with the corresponding time invariant regressors $\bar{\mathbf{x}}_i$.

Let $W = (\omega_1, \mathbf{x}_{it}, \bar{\mathbf{x}}_i, \gamma, \boldsymbol{\beta}, \mathbf{c}, \xi_i)$. In Mundlak's (1978) approach, the likelihood function is (see Appendix [A.1.9](#) for the derivation)⁵⁰

$$f(y_{i2}, \dots, y_{iT} \mid \omega_1, \mathbf{x}_{it}, \bar{\mathbf{x}}_i, \gamma, \boldsymbol{\beta}, \mathbf{c}, \sigma_\xi^2, y_{i1})$$

⁵⁰ We thank Professor Jeffrey Wooldridge (Michigan State University, Department of Economics, USA) for his assistance in the derivation of the likelihood function.

$$= \int_{-\infty}^{\infty} [\prod_{t=2}^T \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{a} + \omega_1 y_{i1} + \xi_i)]^{y_{it}} [1 - \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{a} + \omega_1 y_{i1} + \xi_i)]^{1-y_{it}} d\Phi\left(\frac{\xi_i}{\sigma_\xi}\right).$$

3.4 Average partial effects (APEs)

Our primary goal is to estimate the effects of the explanatory variables $\mathbf{x}_{it}$ on the response probability $P(y = 1|\mathbf{x}, y_{i1}, a_i)$. Wooldridge (2010, p. 277 and pp. 583-584) shows that, for probit, a consistent estimator of the average partial effects (APEs) is

$$(3.38) \quad \hat{\beta}_K [N^{-1} \sum_{i=1}^N g(\mathbf{x}_i \hat{\boldsymbol{\beta}})],$$

when $\mathbf{x}_K$ is continuous, or

$$(3.39)$$

$$N^{-1} \sum_{i=1}^N [G(\hat{\beta}_1 + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_{K-1} x_{i,K-1} + \hat{\beta}_K) - G(\hat{\beta}_1 + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_{K-1} x_{i,K-1})],$$

when $\mathbf{x}_K$ is binary.

Based on Heckman's (1981b) approach, we average out the individual heterogeneity a_i and extend the explanatory variables to include the initial observations y_{i1} (see Wooldridge, 2010, 15.8.4, p. 628). Then, we compute the APEs of all explanatory variables on the response probability $P(y = 1|\mathbf{x}, y_{i1}, a_i)$ by taking derivatives of $\hat{\beta}_K [N^{-1} \sum_{i=1}^N g(\mathbf{x}_i \hat{\boldsymbol{\beta}})]$, with respect to the continuous variables $\mathbf{x}_K$ for each woman, or take differences with respect to discrete variables $\mathbf{x}_K$ for each woman, and then we average these differences across all women, that is, $N^{-1} \sum_{i=1}^N [G(\hat{\beta}_1 + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_{K-1} x_{i,K-1} + \hat{\beta}_K) - G(\hat{\beta}_1 + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_{K-1} x_{i,K-1})]$. This results in estimates of the partial effect of the explanatory variables on the probability of participating in the labour market, averaged across the population of women. In other words, the APE of an explanatory variable is the percentage increase (or decrease) in the probability of participating as a consequence of one unit increase in this variable, after we have controlled for the presence of a_i and the initial conditions y_{i1} .

For the fixed-effects approach of Bartolucci and Nigro (2010), it is not technically feasible to calculate APEs (Lucchetti and Pigini, 2017, top of p. 16).

Chapter 4

The data

4.1 Introduction

This chapter presents the British Household Panel Survey.⁵¹ Its important features, data preparation along with specific issues about health variables are presented in sections 4.2 and 4.3, respectively. Section 4.4 presents an overview of the final dataset with summary statistics,⁵² and section 4.5 summarizes. The empirical results of chapter 5 are based on this dataset.

4.2 The British Household Panel Survey

The British Household Panel Survey (BHPS) is a household-based longitudinal survey of people living in the United Kingdom, carried out by the Institute for Social and Economic Research (ISER) at the University of Essex. The BHPS began in 1991 on a sample of approximately 5,500 households and over 10,000 adult respondents, aged 16 and over in England, Wales, and Scotland. To allow more representative analyses separately for the four British countries, i.e., England, Scotland, Wales, and Northern Ireland, the sample size increased over the years. In 1999, the BHPS included additional samples of about 1,500 households living in each Scotland and Wales and in 2001 a sample

⁵¹Produced by the University of Essex, Institute for Social and Economic Research,(2018). *British Household Panel Survey: Waves 1-18, 1991-2009*. [data collection]. *8th Edition*. UK Data Service. SN:5151, <http://doi.org/10.5255/UKDA-SN-5151-2>

⁵² We used STATA for the derivation of tables and figures of this chapter.

of about 2,000 households living in Northern Ireland was added to the BHPS sample. There are 18 annual waves of the BHPS.⁵³

The first and main purpose of the BHPS is to provide high quality longitudinal data about topics such as labour market, income, education, health, family, social life and, furthermore, to permit interdisciplinary research in many areas such as the relationship between health changes and labour force participation, life cycle variations in income, and the effects of cohabitation and fertility on participation. Hence, the survey helps researchers to identify social and economic changes at the individual and household level and, through estimations, furthers researchers' understanding of the effects of such changes, their causes and consequences, in relation to a range of both objective and subjective indicators that it provides. The final purpose is to offer guidance to authorities to design policy interventions that may impact upon the general well-being of the UK population (<https://www.iser.essex.ac.uk/bhp>).

In order to achieve such high quality measurements of social change in the UK, the BHPS follows the sampled households annually, re-interviewing the same adult members of them in successive waves. The collection of information on changes to the household and individual circumstances is carried out with specific rules. Specifically, interviews are carried out annually with each adult member (aged 16 and over) of the selected households as long as they live in the UK. Nonetheless, individuals may join and leave the sample. The BHPS tracks changes in household formation and dissolution and has a number of following rules that determine who is eligible to be interviewed at each wave. There are three categories: 1) Original Sample Members (OSM). These consist of members of Wave One households, and their natural children born after the start of the study. This group is always eligible to be interviewed. 2) Temporary Sample Members (TSM). These consist of individuals who form households with OSMs after the start of the study. TSMs are eligible to be interviewed for as long as they are living with an OSM, but cease to be eligible if they leave the household. 3) Permanent Sample Members (PSMs) when TSMs

⁵³ Waves correspond to fieldwork periods. For BHPS, the fieldwork period begins at the 1 September for wave one, 1 September for wave two, etc. (British Household Panel Survey User Manual, 2018, volume A, appendix 2, p.236)

have children with OSMs. PSM are eligible to be interviewed even when they stop living with an OSM (British Household Panel Survey User Manual, 2018, volume A, p.25)

From 1997, the BHPS became sub-sample of the European Community Household Panel (ECHP), providing data for the United Kingdom, but it finished in 2008. From 2010 the sample of households from the BHPS has become part of the larger UK household longitudinal survey, called Understanding Society.⁵⁴

We choose to investigate the determinants of female labour supply in the UK from BHPS for the availability of data over a long span along with a large documentation about BHPS itself and a non-negligible documentation in the research questions of our interest. Moreover, we consider that a detailed analysis of one country provides us with more reliable estimates of self-assessed health, because individuals perceive health differently across countries (see, for example, McFadden et al., 2005). Our empirical analysis is based on the first twelve waves of the BHPS, from 1991 to 2002.

4.3 Data preparation and variables of the dataset

4.3.1 Data preparation

There are two record types that we use from the BHPS. Record type HHRESP contains data from the Household Questionnaire for sample selected households and record type INDRESP contains individual data from main individual questionnaire.

In all waves, letters of the alphabet are used as wave indicators. Thus, "a" is attached to all wave 1 variables, "b" to wave 2 variables, and so on. For presentation purposes, we use a "w" attached to variables throughout this chapter, indicating that the variables appear in all waves.

⁵⁴ Therefore, we concentrated only on analysis on waves produced specifically by BHPS data, within the time span BHPS took place. Augmentation of the time span would require combination of the BHPS and Understanding Society datasets and further demanding investigation of the degree of coherence for the variables of our interest between the datasets.

In the BHPS, there are two primary key variables which uniquely identify the household interviewed at the particular wave and any eligible person within the household at a given wave. The first key variable is the household identification number, labeled “wHID” and the second key variable is the person number, labeled “wPNO.” The “reference person” in the household is usually the oldest person within it and normally has a value that equals one (wPNO = 1). Her/his spouse often has a value equal to two (wPNO = 2). Information about relationships between people within households in BHPS is given in terms of these wPNO variables.⁵⁵

In order to merge information for the same respondent in successive waves, we used the record type XWAVEID that contains information for matching individuals between waves using a key variable, the cross-wave person identifier, labeled “wPID”. Hence, for the selection of women for consecutive 12 years, we used XWAVEID, keyed on PID, the Cross-wave person identifier, together with that individual's wHID and wPNO for each year that these are available.

Married women are our target. Therefore, the second stage of constructing the dataset involved identifying women who are continuously married in the BHPS. In each wave, the BHPS gathers information on marital status (wMASTAT variable). The respondents are asked the following question: “Are you currently married, living with a partner, widowed, divorced, separated or never married?” The information is located in wINDRESP files.

The third stage was the construction of a new dataset that combines individual characteristics and household income of married women from wINDRESP and wHHRESP files, respectively. In other words, the individual demographic characteristics of a woman and the monthly household income of her household are observed in a single row for each wave in the final dataset.

In the end, the dataset consists of 332 continuously married women across the 12 waves of the BHPS, observed at wave 1.

⁵⁵Record Type AEGOALT in BHPS provides a mechanism for identifying the relationship of each individual in a household to all others.

4.3.2 Variables of the dataset

The epicenter of our investigation is in demographic characteristics and in health-related variables that may affect a married woman's decision to participate or not. Our analysis is built on the work of Eckstein and Wolpin (1989) who investigated the intertemporal labour supply decisions of married women aged from 39 to 45 years old. Similarly to them, our focus is on married women who are beyond their reproductive age. Hence, the sample is restricted only to women who are between 39 and 45 years old and who are either continuously legally married or cohabitating across all the twelve years of the survey. The time span of the sample permits us to investigate the determinants of labour supply, before women decide whether or not to retire. This special characteristic of our data distinguishes our analysis from the majority of studies that investigate the association between the decision to retirement and health of older individuals and not the association between the decision to participate and health of women of working age.

In each interview, the respondents are asked about their current economic activity. They are asked to select one of the ten labour market states showed in a card for the following question: "Please look at this card and tell me which best describes your current situation? The interviewer shows a card that includes the following possible options: "self-employed," "in paid-employment (full-time or part-time)," "unemployed," "retired," "looking after a family," "full-time student/at school," "long-term sick or disabled," "on maternity leave," "on a government training scheme," "something else," and proxy respondent in case the eligible individual is absent at the time of the interview and another member of the household answers on behalf of her (the variable is labeled as wJBSTAT). We created two categories of employment state for the respondents. One category, named "employed," contains respondents who are self-employed and in paid employment, either full-time or part-time. The second category, named "non-employed," contains respondents who are unemployed and, additionally, the group of respondents who are family carers, full-time students/at school, on maternity leave and those defined as not employed from answers provided by the proxy respondents. This classification is similar to most of the previous literature where it is common either to exclude this group of non-participants in the labour force (also known as economically inactive) or to include them along with the unemployed individuals (see, for example, Campolieti, 2002, Jones, 2006; Garcia-Gomez

et al., 2010). We followed the latter approach and classified this group as non-employed and therefore included them with the unemployed, because we consider them as potential workers in subsequent waves. Retired and long-term sick or disabled respondents are excluded from the analysis because we treat these characteristics as sufficient explanations for their non-participation in economic activity.

Table 4.1 presents a distribution of the dependent variable we focus our interest on, the economic activity spells in the sample in wave 1.

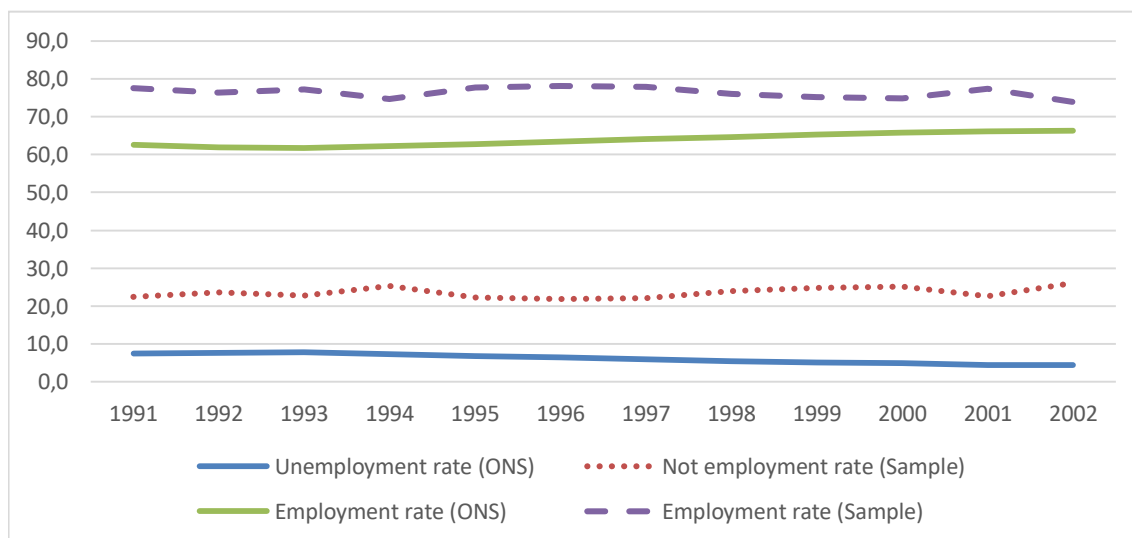
Table 4.1: Current economic activity (Wave 1)

<i>Wave 1</i>	<i>Full sample</i>	<i>Non-employed</i>	<i>Employed</i>
In paid employment	231		231
Self-employed	24		24
Unemployed	4	4	
Maternity leave	9	9	
Family Care	56	56	
Full-time student/school	2	2	
Proxy respondent	3	3	
Government scheme	0		
Missing values	3		
Total	332	74	255

Notice that our sample consists of 332 women. However, we lack information of the economic activity for 3 women in the first wave. This makes a loss of cases less than 1 percent. We consider this percentage a minor loss compared to the total sum of cases (332) and we proceed our analysis without further consideration upon these missing cases in economic activity in the first wave (see Appendix [A.7](#) for further details on missing-values patterns in our final sample).

Figure 4.1 plots the yearly female employment and unemployment rates from the Office for National Statistics (ONS) in the UK and from our dataset derived from the BHPS. The solid lines represent the ONS rates and are calculated from a sample of women aged above 16, irrespectively of their marital status, and the dashed lines represent the BHPS rates calculated from a sample of women aged between 39 and 45 that are either married or cohabiting. All rates are for the period 1991-2002. The pairwise comparisons between the sample rates to the yearly statistics from the ONS demonstrate that sampled married or cohabitating women aged 39-45 have both higher employment and non-employment rates than employment and unemployment rates the general female population of working age has. A justification for higher employment rates from our sample might be due to the age range of our sample. The proportion of older women tends to work more than the younger ones. Difference in unemployment rates are justified from our definition of “non-employed” that includes more groups of women compared to the

Figure 4.1 *Employment and Unemployment – Non-employment rates from the ONS and from the sample*



Notes: Women’s employment and unemployment rates in solid lines are obtained from the ONS. Women’s employment and non-employment rates in dashed lines are our own calculations from the sample.

ONS definition.⁵⁶

Again, our main interest is in the effect of health on labour supply. One measure of health that is broadly used in many studies and is available in the BHPS, is self-assessed health (wHLSTAT) in wINDRESP files. The interviewer asks the respondents to classify their health status according to the question: “Please think back over the last 12 months about how your health has been. Compared to people of your own age, would you say that your health has on the whole been excellent/good/fair/poor/very poor?” Self-assessed health is therefore a subjective measure of health that provides an ordinal five-point scale ranking of health status, relative to individuals’ own concept of health. However, in wave 9 only, question and response categories are modified. The question is now: “In general, would you say your health is: excellent, very good, good, fair, poor?” Note also that the question removes the age effect on health. Hernandez-Quevedo, Jones and Rice (2005) have investigated reporting bias due explicitly to the change in this measure of self-assessed health used in BHPS in wave 9. In particular, by using panel data ordered probit and generalized ordered probit models, they found that collapsing the self-assessed health into four categories does not significantly change the relationship between socio-economic and health characteristics. In order to maximize the time span of data and to achieve consistency over all 12 waves, we followed the method of Hernandez-Quevedo et al. (2005) and recoded self-assessed health as a four-category scale: excellent, good or very good, fair, and poor or very poor. The reference category is “poor or very poor.”

An alternative and a more objective health measure, compared to self-assessed health, is the existence of health limitations (wHLLT) in wINDRESP files. The variable is defined by a response to the question, “Does your health in any way limit your daily activities compared to most people of your age?” This question is more directly related to daily activities and, indirectly, to work. The question is not asked in wave 9. We followed

⁵⁶According to the ONS, unemployed people of working age are those without a job who have been actively seeking work within the last four weeks and are available to start working within the next two weeks. Consequently, the unemployment rate is the proportion of the economically active population (those in work plus those seeking and available to work) who are unemployed. The definition from the ONS meets the International Labour Organization definition of unemployment (<https://www.ons.gov.uk/employmentandlabourmarket>).

Jones et al. (2010) and we assumed that health limitations are likely to consist of chronic problems. Under this consideration, we use the wave 8 values in wave 9.

In order to proceed to a deeper investigation of the association between labour supply and health, we included nine more health-related variables, derived from questions on specific health problems. In each consecutive wave, the respondent is asked whether she has any specific health problems from a list. The health problems are related to arms, legs, hands, etc. (wHLPRBA), to sight (wHLPRBB), to hearing (wHLPRBC), to skin conditions/allergy (wHLPRBD), to chest/breathing (wHLPRBE), to heart/blood pressure (wHLPRBF), to stomach or digestion (wHLPRBG), to diabetes (wHLPRBH) and to anxiety/depression, etc. (wHLPRBI). We constructed a dummy variable for the presence or not of each specific health problem. The preliminary analysis showed strong effects only of health problems related to arms, legs, hands and of health problems related to anxiety/depression (see Appendix [A.2](#) for the correlation matrix). Therefore, from the nine dummy variables, only these two are included in further analysis.

Household income includes labour and non-labour income of all household members. First, we used the Retail Price Index⁵⁷ to adjust the total household income in 1991 sterling pounds values (see Appendix [A.3](#)) and then we divided the resulting income by the McClements Equivalence Scale⁵⁸ (see Appendix [A.4](#)), an equivalence scale after housing costs are deducted⁵⁹ that permits the transition between household income and the individual income in order to make income comparisons among women who come from households with different household size and composition. Theoretically, if we define δ_i as the household income and k_i as the total number of individuals in the household, adults

⁵⁷ In the UK, the Retail Price Index is constructed by the Office for National Statistics (ONS).

⁵⁸ The McClements scale is a globally accepted scale and applies to survey data where the household income is of primary concern among households of varying size and composition (Social Metrics Commission, 2019).

⁵⁹ There are two forms of the McClements scale available, namely before and after cost of housing is deducted. Our analysis is conducted according to the relevant literature with the use of equivalized income after the cost of housing is deducted. That option has two main reasons. First, the cost of housing can vary considerably among women who appear to have otherwise identical circumstances, for instance among women who own their house, pay mortgage or rent a home. Second, the calculation of the equivalized income, after the cost of housing is deducted, is not affected by whether or not housing benefits, which primarily support the poorest households, are considered as wages (British Household Panel Survey User Manual, 2018, volume A).

and children, we obtain the individual income of an equivalent woman as $\widetilde{\delta}_i = \frac{\delta_i}{k_i}$. The McClements scale gives a weight of 0.55 to head, 0.45 to spouse, 0.07 to children aged between 0 and 2 years old, and 0.18 to children between 3 and 4 years old. Clearly, couples with fewer children or couples with children aged less than 2 years old increase the equivalized income (British Household Panel Survey User Manual, 2018, volume A).

Non-labour income enters the model as a determinant of women's labour supply. We followed Hyslop (1999) and created two income variables, the mean (over the twelve years) of non-labour income and the deviation from the mean. The mean income variable helps us to assess the impact of the monthly permanent non-labour income on labour supply decisions, whereas the deviation from the mean helps us to estimate the impact of the monthly transitory non-labour income. In economic theory, permanent income consists of an expected income and transitory income consists of unexpected income. In the permanent income hypothesis, first developed by Friedman (1957),⁶⁰ permanent labour income is expected to have a larger effect on labour supply than transitory labour income. We expect the same effect for our non-labour income variables. In BHPS, monthly permanent non-labour income is derived from monthly household income (wFIHHMN on record wHHRESP) that includes a woman's individual gross earnings from her job, spouse or partner's monthly gross pay and household's other revenues. Then we subtracted the monthly woman's labour income (wFIMNL on record wINDRESP) from monthly household income and we got the non-labour income used in our analysis. All income measures are equivalised and expressed in 1991 sterling-pound values. In our analysis, we used the natural logarithm to measure "percent" changes in permanent and transitory non-labour income.

The presence of children, especially when they are young, is known to have a strong negative effect on the mother's labour supply. The presence of children living in the household at different ages (wNCH02, wNCH34, wNCH511) is included in the analysis as a binary indicator of whether the woman has any children in a specific age group of children or not. The relevant variables are derived from wINDRESP files.

⁶⁰ Friedman (1957) introduced the permanent income hypothesis to describe how economic agents consume over their lifetime. His hypothesis states that agents' consumption at a given time is determined mainly by their expected future income and not by their current income. According to Friedman, changes in transitory income have a smaller effect on consumption than changes in permanent income.

To study the effect of an additional year of education on labour supply decisions, we include years of education in our model from wINDRESP files. In BHPS, education is measured by the highest formal educational qualification (the variable is wQFACHI) completed at each wave (the categories of the variable wQFACHI are HIGHEST DEGREE, 1st DEGREE, HND/Teaching, A LEVEL, O LEVEL/CSE, OTHER QUALIFICATIONS. We converted the educational categories to years of education in order to be in line with Hyslop (1999). Years of education are imputed from the categorical scheme: “HIGHEST DEGREE” = 18 years, “1st DEGREE” = 16 years, “HND/Teaching O/CSE” = 15 years, “A LEVEL” = 13 years, “O LEVEL/CSE” = 11 years. For the last category “OTHER QUALIFICATIONS” we allowed for another variable that indicates school leaving age (wSCEND). If the respondent claims other qualifications and her school leaving age is greater than 16 years, we consider that she has completed 11 years of education. If the respondent claims other qualifications and her school leaving age is less than 16 years, we consider that she has completed the ‘school leaving age’ years minus the lowest compulsory five years of education. We expect positive effect of an additional year of education on the probability of participation.

Female labour supply decisions have also been found to depend strongly on work experience. Work experience represents the total number of years the woman has worked. In BHPS, the employment status history (wBLESLEN) is present at wave 2 only in months. We converted the months to years and for each of the subsequent waves, we added an additional year of work experience if the respondent stated “employed,” otherwise we did not add work experience. We expect a strong positive effect of work experience on female labour supply decisions.

The relationship between a woman's labour supply and her partner's economic activity is also investigated by the inclusion of a variable (wSPJOB) that indicates whether or not the partner is employed. The estimate of the coefficient of this variable measures the effect of spouse's labour market attachment (i.e., usually his job loss) on the probability of woman's labour market attachment and can be interpreted as a measure of the added worker effect. Household labour supply theories predict that a family member would, under certain assumptions, increase his or her labour supply to compensate for the income loss due to the unemployment of the primary breadwinner. Every family member can become an added worker, but empirical research focuses mainly on couples because the

greatest interdependency in household labour supply decisions is found between partners (see Sayli, 2017, p.52, for a review upon the topic). The effect is considered predominantly positive by most studies (see, for example, Del Boca et al., 2000), which means that partner's economic inactivity increases woman's labour market attachment. In the UK, however, there is empirical evidence that the added worker effect might be negative (or reverse). This implies that a woman whose partner is unemployed is more likely to decrease her labour supply, compared to a woman whose partner is employed, unless the couple is in extreme financial distress. In addition, the negative effect is found when the spouse is in long-term unemployment (Harkness and Evans, 2011, as cited in Sayli, 2017, p.58). In her review of literature, Sayli (2017) points out that the negative added worker effect in the UK could be attributed to the potential complementarity of spouses' leisure times (that is, an inactive woman may not work when her spouse gets unemployed because she derives more utility when she spends more time with him), the disincentive effect of the welfare system on women's labour force participation and the social conventions and established division of labour in the household (page 59). Sayli (2017), for example, examines how a cohabitating woman responds to her partner's different labour market activities. Using the BHPS dataset from 1991 to 2009, Sayli (2017) finds that a woman whose partner is unemployed is 23 percent less likely to enter the labour market than a woman whose partner is employed. However, this is also less likely to occur during recessions (Harkness and Evans, 2011; Bryan and Longhi, 2018, as cited in Sayli, 2017, p.58). Notice that, in chapter 5, the dummy variable (wSPJOB) takes on the value of 1 if the spouse or partner is unemployed and the value of zero otherwise.

Finally, to better understand the dynamics of labour supply, national unemployment rate (unempl_T) for the UK is present in the model, for all individuals, men and women, aged 16 and above and for each year from 1991 to 2002. By examining the impact of the unemployment rate on women's labour supply, we investigate the "discouraged worker hypothesis", according to which individuals may stop searching for work and withdraw from the labour market entirely, after failed job search or when facing a discouraged prospect of finding jobs. The "discouraged worker hypothesis" is often empirically confirmed. For example, Xiaodong Gong (2010), in a reduced-form approximation for female labour force participation in Australia, finds that the labour force participation decreases by 1.2 percentage points, for every percentage point increase in the

unemployment rate (page 19). Thus, the “discouraged worker hypothesis” is expected to be negative.

The dataset combines information to create a single file on characteristics and contains information for economic activity, demographic characteristics and health measures for 332 respondents (see Appendix [A.5](#) for summary of definition and construction of variables and Appendix [A.6](#) for summary of record types used from BHPS). This gives a total of 3,984 observations (332 women multiplied by 12 waves).

4.4 Summary Statistics of the dataset

Summary statistics for the variables included in the analysis are presented in table 4.2. The sample is presented for wave 1 and wave 12, separately, and divided by employment state. In the initial year, that is in wave 1, women’s mean age is 42.16 and ranges from 39 to 45 years old, as expected. Average work experience is 17.19 years and average education is 11.52 years. As it can be seen, the majority of women are in employment (77 percent) and 90 percent of them have an employed partner. The majority also reports very good or good health (46 percent). Only 6 percent of women report poor or very poor health. Accordingly, only 9 percent of all women report having health limitations and only 4 percent claim anxiety and depression. It is remarkable that the percentage of women who report poor or very poor health and health limitations is higher for the non-employed, compared to women who are employed (columns 5 and 6). Furthermore, non-employed women have remarkably higher percentages of having children in all children’s age ranges and lower work experience, compared to employed women (columns 5 and 6). In wave 12, we observe the same pattern for health variables. Non-employed women claim higher limiting health problems (45 percent) and report higher percentages of fair and poor or very poor health, compared to employed women. It is notable that non-employed women have on average much lower work experience (18.21 years), compared to employed women (28.14 years) (columns 11 and 12). From wave 1 to wave 12, the mean age of women is now 53.10 years, the mean work experience is 25.51 years and the mean years of education are 11.76 years (column 8). Permanent non-labour income, as expected, remains invariant from wave 1 to wave 12, whereas transitory income almost invariant (column 8).

Table 4.2 Descriptive statistics: Wave 1 and Wave 12

	Wave 1					Wave 12				
	Mean	Min	Max	Mean if y = 0	Mean if y = 1	Mean	Min	Max	Mean if y = 0	Mean if y = 1
Age	42.16	39	45	41.81	42.26	53.10	49	57	52.71	53.20
Employment status	0.77	0	1	0	1	0.74	0	1	0	1
Limiting health problems	0.09	0	1	0.28	0.03	0.17	0	1	0.45	0.07
Health status	1.94	1	4	2.34	1.83	2.19	1	4	2.66	2.03
Excellent health	0.32	0	1	0.23	0.35	0.20	0	1	0.10	0.23
Very good or good health	0.46	0	1	0.39	0.48	0.50	0	1	0.38	0.56
Fair health	0.14	0	1	0.19	0.13	0.18	0	1	0.27	0.13
Poor or very poor health	0.06	0	1	0.19	0.03	0.10	0	1	0.24	0.07
Arms, legs, etc	0.17	0	1	0.29	0.14	0.32	0	1	0.45	0.29
Anxiety/Depression	0.04	0	1	0.09	0.03	0.08	0	1	0.19	0.05
Years of experience	17.19	0	32.33	12.67	18.71	25.51	0	43.33	18.21	28.14
Years of education	11.52	0	18	11.08	11.63	11.76	0	18	11.08	12.00
Presence of children aged 0-2 years	0.02	0	1	0.08	0.00	0.00	0	1	0.00	0.00
Presence of children aged 3-4 years	0.05	0	1	0.19	0.02	0.00	0	1	0.01	0.00
Presence of children aged 5-11 years	0.33	0	1	0.55	0.27	0.04	0	1	0.06	0.03
Transitory Income	0.00	-5.50	2.34	0.00	-0.12	0.02	-6.25	2.28	0.12	0.00
Permanent Income	6.99	4.82	8.37	6.87	7.02	6.99	4.82	8.37	6.96	7.00
Spouse employed	0.90	0	1	0.78	0.94	0.78	0	1	0.59	0.88
Unemployment rate	8.90	8.90	8.90	8.90	8.90	5.20	5.20	5.20	5.20	5.20

Note: y = 0 refers to the non-employed women and y = 1 refers to employed women.

Table 4.3 presents the variables used in the analysis along with their respective summary statistics for the full sample, across all 12 waves. As expected, the women's age ranges from 39 to 57 years; If she is at most 45 years old at wave 1, she is 57 at wave 12. Women, on average, are aged about 47 years old, have 21.39 years of work experience and 11.65 years of education. The majority of women in our sample work (76 percent) and cohabitate with an employed partner (85 percent). The percentage of women reporting health limitations is 14 percent in the total sample and is significantly greater for the non-employed women (35 percent) compared to the employed (7 percent). Additionally, on average, 25 percent of women report health problems related to arms, legs, hands, etc. and 7 percent related to anxiety and depression. Interestingly, the percentage of women who report excellent health rises from 13 percent for the non-employed to 24 percent for the employed and, the percentage of women who report poor or very poor health falls from 21 percent for the non-employed to 4 percent for employed women.

Table 4.3 Descriptive statistics: all waves

	Mean	Min	Max	Mean if y = 0	Mean if y = 1
Age	47.61	39	57	47.24	47.65
Employment status	0.76	0	1	0	1
Limiting health problems	0.14	0	1	0.35	0.07
Health status	2.13	1	4	2.54	2.00
Excellent health	0.22	0	1	0.13	0.24
Very good or good health	0.50	0	1	0.39	0.54
Fair health	0.19	0	1	0.24	0.17
Poor or very poor health	0.08	0	1	0.21	0.04
Arms, legs, hands	0.25	0	1	0.37	0.21
Anxiety/Depression	0.07	0	1	0.17	0.05
Years of experience	21.39	0	43.33	15.61	23.14
Years of education	11.65	0	18	10.98	11.86
Presence of children aged 0-2 years	0.02	0	1	0.05	0.00
Presence of children aged 3-4 years	0.02	0	1	0.06	0.00
Presence of children aged 5-11 years	0.15	0	1	0.28	0.12
Transitory Income	0.00	-6.67	2.59	0.02	0.00
Permanent Income	6.99	4.82	8.37	6.97	7.00
Spouse employed	0.85	0	1	0.67	0.92
Unemployment rate	7.52	5.10	10.40	7.50	7.55

Note: y = 0 refers to the non-employed women and y = 1 refers to employed women

Non-employed women in the full sample have more children than employed women and more children than average. The difference in percentages is very apparent for the presence of children aged 5-11 years. All percentages indicate the commonly held consideration that the presence of children tends to reduce women's labour supply.

The longitudinal data provided by the BHPS allows us analysis of transitions over time. Given the emphasis on dynamics of the labour supply decisions, we begin with a transition matrix of the employment states. The rows of the standard two-way table 4.4 indicate current employment state and the columns show employment states at the previous year. The high percentages on the diagonal of the matrix show a strong degree of persistence. That is, women who did not work at the previous year and do not work at the current year are 86.36 percent and women who worked at the previous year and also work at the current year are 95.62 percent. These percentages give an indication that the persistence may be due to true state dependence. The corresponding Pearson χ^2 test for the independence of the rows and columns informs us that there is strong association between employment states at current and previous years, namely $\chi^2(1) = 2,607.28$ ($p - value < 0.000$). The test indicates that employment state of previous year is strongly predictive for the work state of current year. The observed differences are significant.

Table 4.4 Employment state transitions

	Non-employed at previous year	Employed at previous year	Total
Non-employed at current year	709	112	821
	86.36	13.64	100.00
Employed at current year	120	2.622	2.742
	4.38	95.62	100.00

In the next two tables, we tabulate employment states by reported self-assessed health status and health limitations. Table 4.5 shows that of those who report excellent health and of those who report good or very good health, the vast majority are in

employment (85.28 and 81.65 percent, respectively). Among women who report poor or very poor health, 62.85 percent are non-employed.

Table 4.5 *Employment state by self-assessed health*

y	Health over last 12 months				Total
	Excellent health	Very good or good health	Fair health	Poor or very poor health	
Non-employed at current year	127	366	228	203	924
	14.72	18.35	31.02	62.85	23.60
Employed at current year	736	1,629	507	120	2,992
	85.28	81.65	68.98	37.15	76.40
Total	863	1,995	735	15	3,916
	100.00	100.00	100.00	100.00	100.00

Table 4.6 tabulates employment states by health limitations. It is clear that of those women who report health limitations, the majority of them are not in employment (61.01 percent).

Table 4.6 *Employment state by health limitations*

y	Health limitations		Total
	No	Yes	
Non-employed at current year	597	327	924
	17.65	61.01	23.58
Employed at current year	2,786	209	2,995
	82.35	38.99	76.42
Total	3,383	536	3,919
	100.00	100.00	100.00

Tables 4.7, 4.8 4.9 and 4.10 summarize transitions across the 12 waves. Table 4.7 shows that the transition of employment states, whether women are employed or non-employed, remains almost invariant across time. Table 4.8 depicts the self-assessed health

transitions. Women who report excellent health gradually decrease from wave 1 to wave 12 by about 37 percent, from 32.63 percent to 20.48 percent, whereas women who report poor or very poor health increase by about 71 percent. Similarly, health limitations transitions in table 4.9 vary across time. We observe a gradual small decrease in women who do not report health limitations and rather a gradual large increase in women who report health limitations. By wave 12, 57 women claim health limitations. Compared to 29 women in wave 1, this is a nearly one hundred percent increase. Finally, table 4.10 shows that the vast majority of women report that they do not have health problems related to anxiety, depression, etc. This percentage is above 90 percent in almost every wave and stays almost invariant across all waves. The percentage of them who report anxiety and depression, however, increases across time. From wave 1 to wave 12, this percentage augments from 4.22 to 8.43 percentage points.

Table 4.7 Employment states by wave

	1	2	3	4	5	6	7	8	9	10	11	12	Total
Non-employed ($y = 0$)	74	78	75	84	74	72	73	79	80	81	72	83	925
	22.49	23.56	22.73	25.38	22.36	21.88	22.19	23.94	24.77	25.16	22.64	26.10	23.59
Employed ($y = 1$)	255	253	255	247	257	257	256	251	243	241	246	235	2.996
	77.51	76.44	77.27	74.62	77.64	78.12	77.81	76.06	75.23	74.84	77.36	73.90	76.41

Table 4.8 Health status by wave

	1	2	3	4	5	6	7	8	9	10	11	12	Total
Excellent health	108	93	82	88	73	66	72	66	34	57	69	68	876
	32.63	28.01	24.70	26.59	21.99	19.88	21.69	19.94	10.30	17.17	20.78	20.48	22.02
Good or very good health	154	165	166	151	166	160	153	168	224	165	176	169	2.017
	46.53	49.70	50.00	45.62	50.00	48.19	46.08	50.76	67.88	49.70	53.01	50.90	50.69
Fair health	48	46	58	65	70	76	75	68	56	80	54	59	755
	14.50	13.86	17.47	19.64	21.08	22.89	22.59	20.54	16.97	24.10	16.27	17.77	18.97
Poor or very poor health	21	28	26	27	23	30	32	29	16	30	33	36	331
	6.34	8.43	7.83	8.16	6.93	9.04	9.64	8.76	4.85	9.04	9.94	10.84	8.32

Table 4.9 Health Limitations by wave

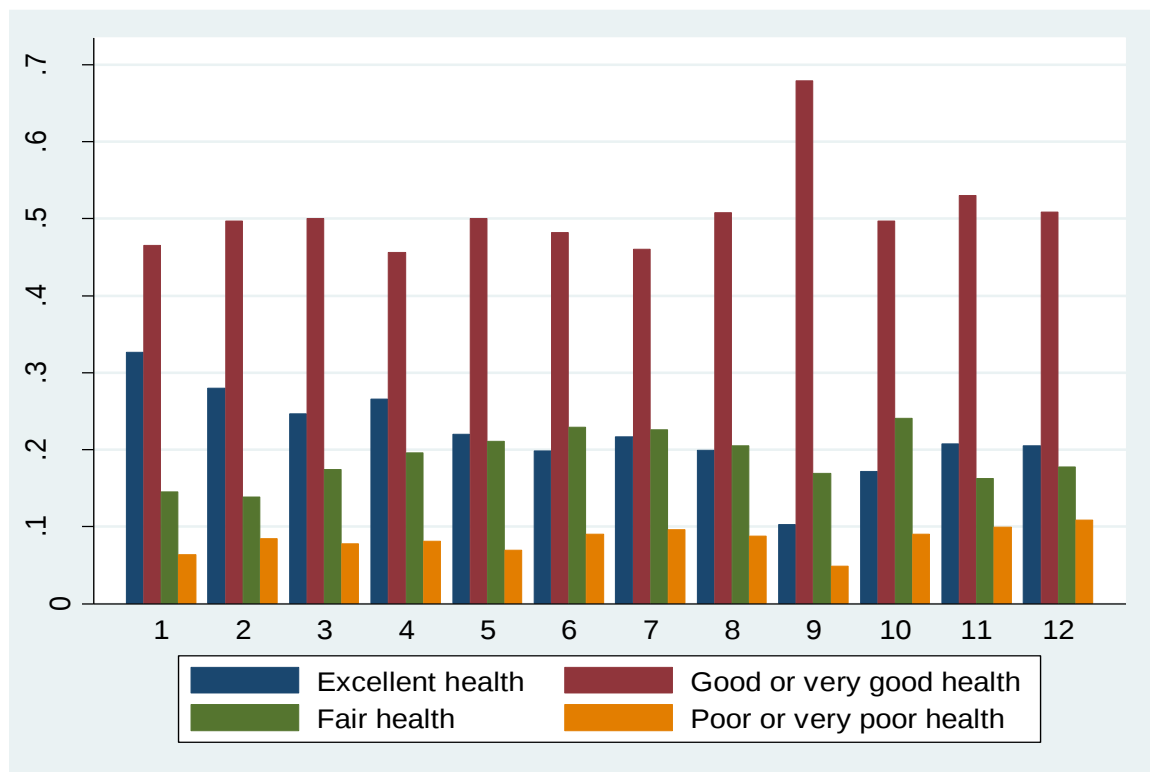
	1	2	3	4	5	6	7	8	9	10	11	12	Total
No (HLLT_y = 0)	303	296	293	292	288	284	283	282	282	277	275	275	3,430
	91.27	89.16	88.52	88.22	86.75	85.54	85.24	84.94	84.94	83.43	82.83	82.83	86.14
Yes (HLLT_y = 1)	29	36	38	39	44	48	49	50	50	55	57	57	552
	8.73	10.84	11.48	11.78	13.25	14.46	14.76	15.06	15.06	16.57	17.17	17.17	13.86

Table 4.10 Anxiety/Depression by wave

	1	2	3	4	5	6	7	8	9	10	11	12	Total
No	318	310	308	309	307	304	302	307	300	299	296	304	3,664
	95.78	93.66	93.62	93.35	92.47	91.57	90.96	92.75	90.36	90.06	89.16	91.57	92.11
Yes	14	21	21	22	25	28	30	24	32	33	36	28	314
	4.22	6.34	6.38	6.65	7.53	8.43	9.04	7.25	9.64	9.94	10.84	8.43	7.89

Figure 4.2 depicts the shape of the distribution of self-assessed health for each wave separately. In all 12 waves, the modal value is good or very good health. The distribution is asymmetric and skewed to the right. The tail of the distribution on the positive side represents the percentage of women who report poor or very poor health. Over time, we can notice a decrease in women who report excellent health and an increase in women who report poor or very poor health. Also notice that the distribution of self-reported health is different at wave 9 than at waves 1–8 or waves 10–12 and it is due to rewording of the self-assessed health variable in the BHPS only at wave 9.⁶¹

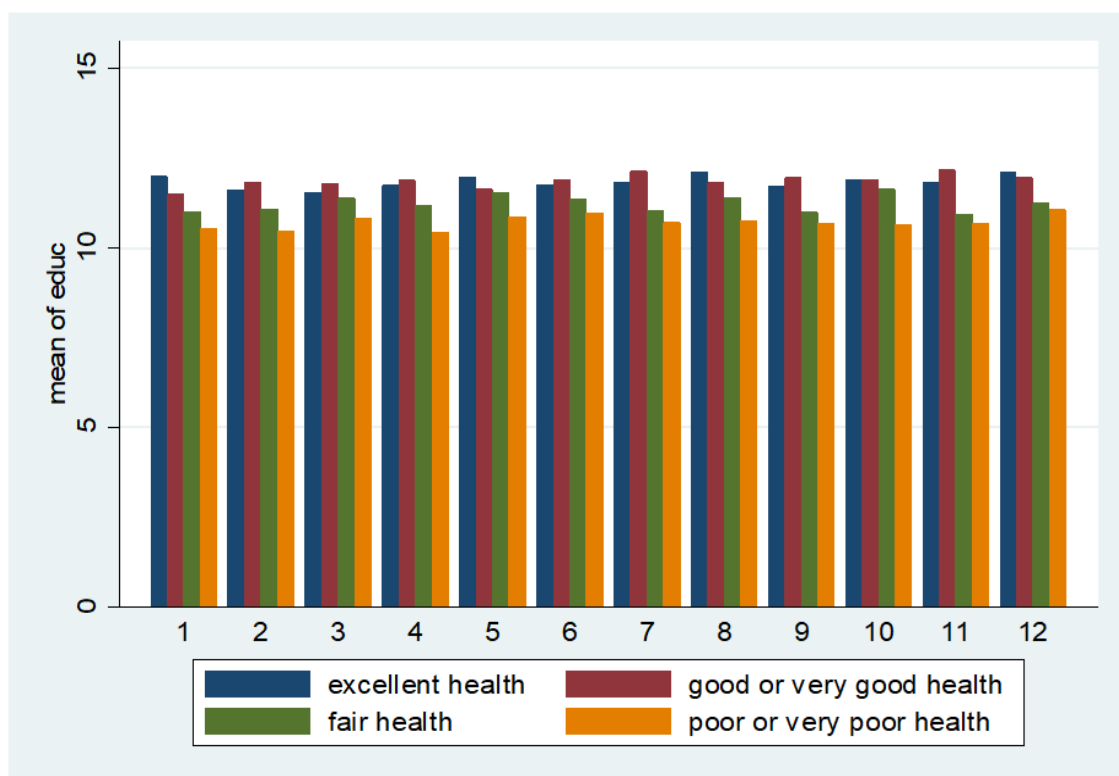
Figure 4.2 *Self-assessed health by wave*



⁶¹ At waves 1–8 and waves 10–12, the respondents are asked the question: “Compared to people of your own age, would you say your health over the last 12 months on the whole has been: excellent, good, fair, poor, very poor?”, whereas at wave 9, the respondents are asked the question: “In general, would you say your health is: excellent, very good, good, fair, poor?”

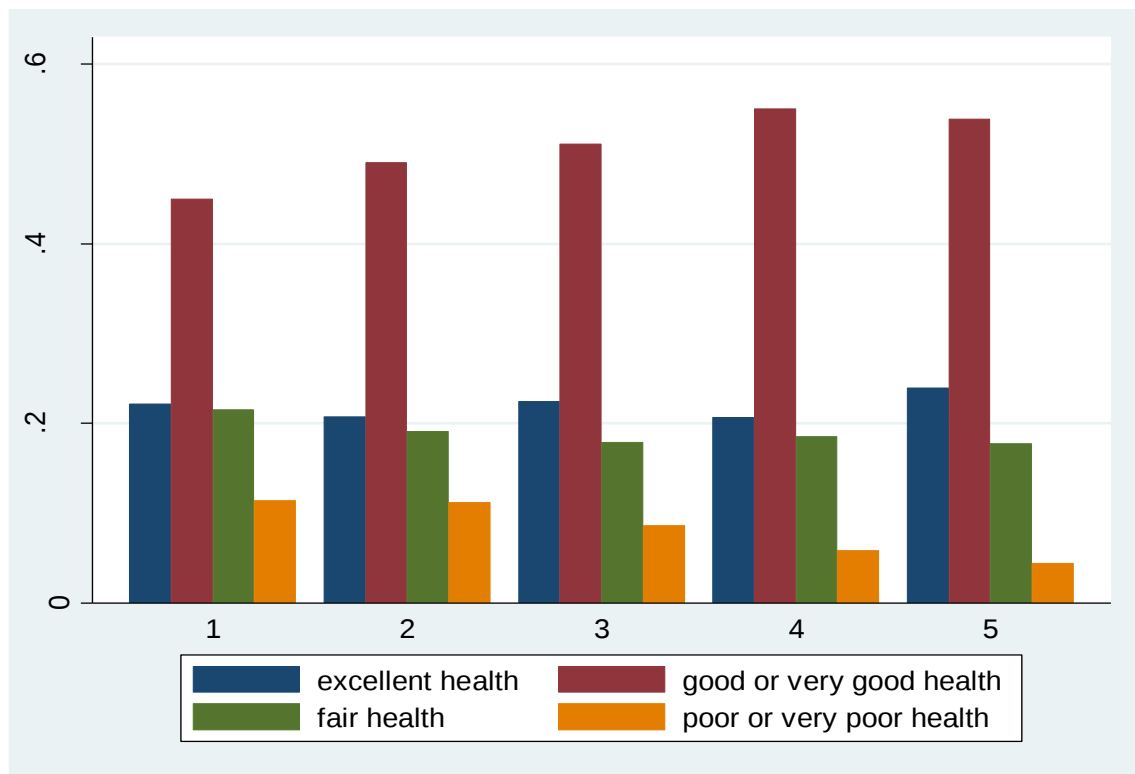
Figure 4.3 depicts the shape of the distribution of years of education split by levels of self-assessed health for each wave separately. It becomes apparent that more years of education are associated with reporting excellent and good or very good health and fewer years of education with reporting fair and poor or very poor health. The distribution is asymmetric and skewed to the right across all waves, with the tail of the distribution representing the percentage of women who report poor or very poor health.

Figure 4.3 Education by self-assessed health by wave



The diagram in figure 4.4 explores the association between health states and monthly non-labour income. The non-labour income is divided into 5 quintiles of the distribution of income, over the 12 waves of the data. The figure shows that there is a notable association between health states and monthly non-labour income. When we move from the poorest quintile (labeled as “1”) to the richest (labeled as “5”), we can observe an increase in the proportion of women reporting good or very good health and a decline in the proportion of women reporting poor or very poor health.

Figure 4.4 *Self-assessed health by quintiles of non-labour income*



4.5 Summary/Overview

Compared with cross-section data, the longitudinal data gave us the opportunity to analyze transitions over time because we have repeated observations for the same individual over time. Tables and graphs demonstrated strong associations between employment states (equivalently, labour supply decisions) and socio-demographic variables. In particular, they revealed a large effect of health variables on labour supply decisions.

Another advantage of the longitudinal data is that it encourages econometric techniques such as random and fixed effects methods. These techniques allow us to control for individual-specific time-invariant characteristics that are not observed in the dataset, referred to as “individual unobserved heterogeneity.” Such characteristics may influence the work decision. In our analysis, we observed that women who claim excellent health are more likely to be employed than those who claim poor health. Consequently, there

might be individual-specific time-invariant characteristics, such as high levels of inner motivation and taste habits in favour of work, that explain this persistence to work. Thus, any correlation between labour supply and health may simply reflect differences in labour supply between, for example, highly motivated and less motivated women. If individual unobserved heterogeneity stays invariant over time, random and fixed effects methods can be used to identify such causal effects. Along with individual time-invariant heterogeneity, we attempt to find the presence and magnitude of true state dependence and serially correlated error terms. In this attempt, we specify dynamic probit models and a dynamic logit model. These methods are analyzed in the next chapter.

Chapter 5

Estimation Results

5.1 Introduction

Our primary interest is in the dynamics of women's participation in the labour market and in the effect of health-related variables on labour force participation decisions over a period of twelve years. To explore whether and to what extent persistence (that is, true state dependence γ , time-invariant individual heterogeneity α_i , and serially correlated error terms ε_{it}), is present and affects our estimates, we specify a simple dynamic pooled probit, a standard dynamic random effects probit, a static random effects probit and three dynamic panel probit models specified by Heckman (1981b), Hyslop (1999), and Keane and Sauer (2009). We also employ the approach of Bartolucci and Nigro (2010), in which the unobserved individual heterogeneity α_i is eliminated by conditioning on the total score. Additionally, we attempt to examine the endogeneity issue of the health-related variables and potential degree of bias in the estimate of state dependence by employing a random effects model that allows the unobserved individual heterogeneity α_i to be correlated with the time means of the time-varying explanatory variables x_{it} , as specified by Mundlak (1978)⁶², embedded in Hyslop's (1999) approach.

In all models, we set y_{it} as the dependent binary variable equal to 1 if a woman participates in the labour market and equal to 0 if she does not participate.⁶³ In an attempt to capture state dependence in participation decisions, we include lagged participation (y_{it-1}) among the regressors. The estimate of the coefficient of y_{it-1} is of primary interest and measures the plausible effect of previous participation state (or labour force participation decision) on the current one.

⁶² See, for example, Stewart (2007), Arulampalam and Stewart (2009), Oguzoglu (2010) and Oguzoglu (2016) who perform Mundlak's (1978) approach in their analyses.

⁶³ In this chapter, whether a woman chooses to participate in the labour market or not is identified by the two categories of employment state for the respondents ("employed," and "non-employed") in chapter 4.

We define $\mathbf{x}_{it}$ as the set of explanatory variables, which may be associated with the response variable, y_{it} . These are a dummy variable for the existence of health limitations, which takes on the value of 1 if the respondent has a health limitation, and the value of zero otherwise; three dummy variables for the reported self-assessed health (the baseline category is the “poor or very poor health” dummy variable); a dummy variable for health problems related to “arms, legs, hands, etc.,” which takes on the value of 1 if the respondent has a health problem with her “arms, legs, hands, etc.” and the value of zero otherwise; a dummy variable for health problems related to “anxiety, depression, etc.,” which takes on the value of 1 if the respondent has a health problem with “anxiety, depression, etc.” and the value of zero otherwise; years of work experience; years of education; dummy variables for the presence of children in the household in successive age groups; variables for monthly non-labour transitory and permanent income; a dummy variable whether the partner/spouse is unemployed; and the national unemployment rate. A woman’s own wage is not included in the analysis. There are at least two reasons for this. First, wages are not observed for women who do not work. Second, in case they were observed, they might be endogenous to labour supply (Cai, 2018, Hyslop, 1999). This approximation, without women’s wage among predictors, is classified as a reduced-form labour supply model (Killingsworth, 1983).

The econometric analysis in this chapter is implemented with the statistical software Gretl and, in particular, with the DPB (dynamic panel binary) function package (Lucchetti and Pignini, 2017), except for the static random effects probit model in table 5.1, which is implemented with STATA.

5.2 Results and Discussion

We present the results of our estimation according to all approaches described in chapter 3.⁶⁴ We begin with the results from the simple dynamic pooled probit and the

⁶⁴ In chapter 4, our descriptive statistics were applied for 332 women. In the present econometric analysis, the sample reduces to 331 women because there are missing observations for the variable “years of education” from one woman for all her waves. Gretl and Stata drop this woman from all probit and logit models analysis.

standard dynamic random effects probit model.⁶⁵ We use these two dynamic nonlinear models on purpose to contrast them with the nonlinear random effects models proposed by Heckman (1981b), Hyslop (1999), and Keane and Sauer (2009). In what follows, the statistical assumptions are progressively relaxed, moving from the simple dynamic pooled probit to the Keane and Sauer (2009) approach, and estimation of additional parameters is needed.

The simple dynamic pooled probit model, assumes that the individual heterogeneity a_i that causes the composite error term, $v_{it} = a_i + \varepsilon_{it}$, to be autocorrelated, is not present ($var(a_i) = 0$). The model imposes that all individuals share the same regression intercept.

Table 5.1 presents the coefficient estimates for the simple dynamic pooled probit and the standard dynamic random effects probit models with their corresponding asymptotic standard errors.

Column [1] gives the simple dynamic pooled probit estimates and the related robust standard errors. All the health variables are of the expected sign and highly significant, except for the variable related to problems with arms, legs, hands, etc. Limiting health problems and health problems related to anxiety and depression reduce the probability of participation, whereas excellent self-assessed health increases the probability of participation. The presence of children in all three age groups reduces the probability of participation. In particular, the younger the children are, the more the

⁶⁵ The terms are defined below.

Table 5.1 Static and Dynamic probit models

	Simple Dynamic Pooled Probit [1]	Standard Dynamic RE Probit [2]	Static RE Probit [3]
Constant	-0.5733 (0.6056)	-0.4515 (0.8947)	-0.0460 (1.5863)
Participation lagged	2.3263*** (0.08188)	2.0167*** (0.1126)	- -
Limiting Health Problems	-0.3780*** (0.1219)	-0.4419*** (0.1432)	-0.7762*** (0.1458)
Excellent Health	0.6721*** (0.1702)	0.7456*** (0.2015)	0.5388*** (0.2043)
Very Good/Good Health	0.5032*** (0.1522)	0.5854*** (0.1789)	0.3992** (0.1814)
Fair Health	0.4176*** (0.1514)	0.5004*** (0.1750)	0.3562** (0.1717)
Arms, legs, hands, etc.	0.0088 (0.0958)	-0.0064 (0.1177)	-0.1441 (0.1241)
Anxiety/Depression	-0.2775** (0.1313)	-0.3825** (0.1640)	-0.4207** (0.1724)
Years of Experience	0.0338*** (0.0054)	0.0466*** (0.0083)	0.0935*** (0.0126)
Years of Education	0.0699*** (0.0171)	0.0998*** (0.0263)	0.1857*** (0.0414)
Presence of children aged 0-2 years	-0.5365** (0.2384)	-0.7270*** (0.2747)	-1.4963*** (0.3001)
Presence of children aged 3-4 years	-0.4549* (0.2487)	-0.6180** (0.2779)	-1.2072*** (0.2654)
Presence of children aged 5-11years	-0.3068*** (0.0975)	-0.3799*** (0.1225)	-0.4991*** (0.1266)
Transitory income	-0.0212 (0.0712)	-0.0379 (0.0795)	-0.2545*** (0.0781)
Permanent income	-0.4395*** (0.0875)	-0.5627*** (0.1328)	-0.8430*** (0.2333)
Spouse unemployed	-0.1465 (0.1068)	-0.1791 (0.1127)	-0.8996*** (0.1740)
Unemployment rate	0.04412** (0.0207)	0.0578** (0.0237)	0.1321*** (0.0258)
lnsigma2		-1.0576*** (0.3276)	0.8090*** (0.1601)
n	3512	3512	3865
Sigma	-	0.5892	1.4985
rho	-	0.2577	0.6919
Pseudo-R ²	0.616	0.622	0.478
lnL	-711.30	-699.73	-973.71

Notes: 1. * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. Standard errors are in parentheses. 2. The model includes time dummies (waves), but we do not incorporate them in the present table as they were not significant. The same applies to the subsequent models.

probability of participation decreases. Education and work experience have positive effects on the probability of participation and are both significant at the 1% level. Notice that the estimate of non-labour permanent income is significantly negative at the 1% level, whereas the estimate for non-labour transitory income is also negative but insignificant. Whether the spouse is unemployed reduces the probability of participation but it is not statistically significant. The estimate of the coefficient of the unemployment rate is significant at the 5% level, but positive, which does not support the “discouraged worker hypothesis.” Finally, participation at the previous year $t - 1$ raises the probability of participation at next year t significantly at the 1% level. This is the estimate of state dependence γ .

The likelihood ratio test statistic with 17 degrees of freedom is 2,380.2 with p-value < 0.001 . The null hypothesis is that all coefficients, other than the constant, are jointly zero, and this hypothesis is strongly rejected, even at levels below 0.001. For pooled probit, R^2 is McFadden’s pseudo- R^2 and equals 0.616.⁶⁶ The log-likelihood of this model is -711.30.

Colum [2] gives the standard dynamic random effects probit estimates. We treat lagged participation as contemporaneously exogenous and we introduce a distinguishable error component for individual heterogeneity a_i , as in equation (3.15). The decomposition of the composite error v_{it} , can provide more efficient estimates and provide information on the extent of the random variability in participation that is due primarily to the individual heterogeneity a_i . The key assumptions for the composite error are homoscedasticity and no serial correlation. That is, $\sigma_v^2 = \text{var}(v_{it}) = \text{var}(a_i + \varepsilon_{it}) = \sigma_a^2 + \sigma_\varepsilon^2$ and $\text{corr}(v_{it}, v_{is}) = E(v_{it} v_{is}) = E(a_i + \varepsilon_{it})(a_i + \varepsilon_{is}) = E(a_i^2) = \sigma_a^2$, for $t \neq s$. We assume additionally that the individual heterogeneity a_i is uncorrelated with the observed regressors x_{it} . Notice that in order to make the parameters of the dynamic random effects probit models identifiable, since we do not have a natural scale for the latent variable in equation (3.15), we need a normalization. The normalization is required because the composite error v_{it} contains two unobservable components, a_i and ε_{it} ; in

⁶⁶ In this thesis we use the adjusted formula for McFadden’s pseudo- R^2 , $1 - (\mathcal{L}_{ur} - k)/\mathcal{L}_0$, where $\mathcal{L}_{ur}$ is the log-likelihood function for the model we estimate, $\mathcal{L}_0$ is the log-likelihood function for the model with only an intercept (and no predictors), and k is the number of explanatory variables in the model. McFadden’s pseudo- R^2 is adjusted for the number of parameters to estimate. It is called “pseudo” as it is treated as an analogy to the R^2 value in the context of linear regression, but it does not reflect proportion of variation in the dependent variable explained by the predictors.

terms of an equation with two unknown terms, we need to set a value in one of them, so that the equation can be solved. A convenient normalization is the variance of the idiosyncratic error term ε_{it} to be restricted to equal 1, for probit models. Therefore, because $\sigma_v^2 = \sigma_a^2 + \sigma_\varepsilon^2$ and we normalize σ_ε^2 to be equal to 1, we can estimate σ_a^2 and hence, we can calculate the proportion of the total unexplained variation in participation that is attributed to a_i , as $\sigma_a^2/(\sigma_a^2 + \sigma_\varepsilon^2) = \sigma_a^2/(\sigma_a^2 + 1)$. The proportion $\sigma_a^2/(\sigma_a^2 + \sigma_\varepsilon^2)$ is called the intra-group (or intraclass) correlation coefficient, *rho*, and indicates the contribution of individual heterogeneity, divided by the variance of the composite error. Theoretically, the intra-group correlation coefficient may have values from close to zero (σ_ε^2 is much larger than σ_a^2) to near one (σ_ε^2 is smaller than σ_a^2) (Liljequist, D., Elfving, B. and Skavberg Roaldsen, K., 2019, p. 6).

The estimates of all coefficients in the standard dynamic random effects probit model are larger (in absolute value) for all variables, compared to the estimates of the simple dynamic pooled probit model (with the exception of the coefficient of lagged participation and the coefficient of health variable related to problems with arms, legs, hands, etc.). The same applies for the standard errors. Notice that the estimate of the coefficient for state dependence in the dynamic random effects probit model is reduced by about 13.5% compared to its counterpart in the dynamic pooled probit, but remains highly significant at the 1% level. This result is in accordance with the finding of Stewart (2007, page 515), who concludes that “if the unobserved heterogeneity exhibits persistence over time, then ignoring it will lead to an overstatement of the true state dependence in unemployment.” This finding is also in line with Stewart (2006, page 266). Specifically, for the health-related variables, we notice that in the standard dynamic random effects probit model the coefficients on “excellent,” “very good” or good” and “fair” self-assessed health variables increase, and the coefficient on “limiting health problems” increase (in absolute value), compared to the pooled probit estimates. Finally, the estimate of the coefficient of the presence of children aged 0-2 years old becomes significant at the 1% level; the estimated coefficient of the presence of children aged 3-4 years old becomes significant at the 5% level; and the health variable related to problems with arms, legs, hands, etc., gets the expected negative sign, but remains insignificant. Similarly to the simple dynamic pooled probit, the estimate of the coefficient of whether the spouse is unemployed is negative, indicating a negative added worker effect, but insignificant.

Finally, the unemployment rate is positive, indicating that there is no evidence of the “discouraged worker hypothesis”.

Nevertheless, as Stewart stresses (2006, page 264; 2007, page 520), care should be taken with comparisons between the estimates of column [1] and [2] because the simple dynamic pooled probit model and the standard dynamic random effects probit model involve different normalizations. The simple dynamic pooled probit estimator uses $\sigma_v^2 = 1$, where $v_{it} = \alpha_i + \varepsilon_{it}$ forms the composite error, whereas the standard dynamic random effects probit estimator uses a normalization of $\sigma_\varepsilon^2 = 1$. Therefore, for comparison, the standard random effects probit model estimates need to be multiplied by an estimate of $\sigma_\varepsilon/\sigma_v = \sqrt{1 - rho}$, where $rho = corr(v_{is}, v_{it}) = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\varepsilon^2)$, where $t, s = 2, \dots, T; t \neq s$. The estimated value of the intra-group correlation coefficient, labeled as rho in the table, is equal to 0.2577 [= $0.589^2/(0.589^2 + 1)$]. Hence, in order to get the appropriate scaled coefficient estimates for the dynamic random effects probit model, we multiply the estimates in column [2] by $\sqrt{(1 - 0.2577)} = \sqrt{0.7423} = 0.8615$. Allowing for the different normalizations, the scaled coefficient estimate on lagged participation is 1.7373, reduced further by 25% if compared with the pooled probit estimate. The same applies for the other estimates.

For the standard dynamic random effects probit model, Table 5.1 also includes the additional individual-level variance component, parameterized as the log of the standard deviation σ_α , labeled Lnsigma2. The standard deviation of the individual heterogeneity σ_α is also included in the table, labeled Sigma, and is equal to 0.5892. The estimated value of the intra-group correlation coefficient, labeled as rho, implies that 25 per cent of the total unexplained variation in participation is attributed to the individual heterogeneity α_i , which suggests a non-negligible degree of persistence due to individual heterogeneity.

We can also perform a likelihood ratio statistic that has an approximate χ^2 distribution under the null hypothesis, with degrees of freedom equal to the number of restrictions being tested (Wooldridge, 2009, p.580).⁶⁷ In this case, the additional parameter (the restriction) is σ_α^2 , the null hypothesis is that σ_α^2 is zero, and the χ^2 test statistic has

⁶⁷ The likelihood ratio statistic is twice the difference in the log-likelihoods, $LR = 2(\mathcal{L}_{ur} - \mathcal{L}_0)$, where $\mathcal{L}_{ur}$ is the log-likelihood function for the model we estimate and $\mathcal{L}_0$ is the log-likelihood function for the model with only an intercept and no predictors associated with it (Wooldridge, 2009, p.580).

one degree of freedom. The gain of fit in model [2] is highly significant, as $\chi^2(1) = 23.2$ (p -value < 0.001), so the null hypothesis ($\sigma_\alpha^2 = 0$) is rejected. Thus, the intra-group correlation coefficient is different from zero, and individual heterogeneity is present. Finally, for the standard dynamic random effects probit model, McFadden's pseudo- R^2 equals 0.622 and the log-likelihood of this model is -699.73.

For comparison, we also include the static random effects probit model for which the coefficient γ in equation (3.15) is set to zero (column [3]). This means that the model ignores possible dynamic effects of the previous participation outcome on the current participation decision. Likewise, for the standard dynamic random effects probit model, the composite error v_{it} is decomposed into two components, $v_{it} = a_i + \varepsilon_{it}$, where the first component a_i is time-invariant individual heterogeneity, which is assumed to be distributed as $N(0, \sigma_\alpha^2)$, and which generates serial correlation in the composite error, v_{it} . The second component, ε_{it} , is assumed to be serially uncorrelated, conditionally independent of a_i and distributed as $N(0, \sigma_\varepsilon^2)$.

Column [3] contains the results from the static random effects probit model. The model is estimated by maximum likelihood. The results show that the static random effects model overestimates the standard deviation of the individual heterogeneity, σ_α . From the definition of rho given above, this model shows that about 69 percent of the latent error variance can be attributed to unobserved heterogeneity. Comparing this estimate of rho with that obtained from the standard dynamic random effects probit model indicates that the extent of individual heterogeneity is dramatically overstated in a model that neglects possible dynamic effects of previous participation outcomes on current participation by a factor of almost 3 (0.6919 versus 0.2577). However, in this approach, all other variables exhibit the expected sign and are statistically significant, except for the health variable related to problems with arms, legs, hands, etc. which has the expected negative sign, but remains statistically insignificant, and except for the effect of the unemployment rate, which is positive, but significant at the 1% level. The variable that indicates whether the partner is unemployed is significant at the 1% level, in this approach. Its sign signifies a negative added worker effect⁶⁸ and implies that a woman whose partner is unemployed is

⁶⁸ The majority of empirical research characterize the added worker effect as the labour supply response of a woman to her spouse's loss of job involuntarily.

more likely to be non-participant than a woman whose partner is employed. Our finding verifies empirical studies, which find that the added worker effect is negative in the UK.

Finally, for the static random effects probit model McFadden's pseudo- R^2 equals 0.478 and the log-likelihood of this model is -973.71

To this point, we have explored three approaches. The first approach is the simple dynamic pooled probit model, in which we assume no individual heterogeneity a_i in our model. The second approach is the standard dynamic random effects probit model, which assumes individual heterogeneity a_i and contemporaneous exogeneity of the lagged dependent variable y_{it-1} . The significance of individual heterogeneity a_i in the second approach suggests misspecification of the simple pooled probit model and likely biased estimates. Therefore, we are in favor of the standard dynamic random effects probit model. For reference, we also implemented the static random effects probit model, which takes into account individual heterogeneity, but ignores state dependence, thus overstating the impact of individual heterogeneity. Also notice that the number of observations is diminished when dynamic models are performed, because in this case the observations of the first wave do not participate in the analysis.

In terms of goodness-of-fit, McFadden's pseudo- R squareds, from the two dynamic approaches, are almost identical (in the simple dynamic pooled probit model, pseudo- R^2 equals 0.616 and in the standard dynamic random effects probit model, pseudo- R^2 equals 0.622) and indicate a good fit.⁶⁹

Table 5.2 below illustrates a comparison of the simple dynamic pooled probit and the standard dynamic random effects probit model in reference to correct predictions of actual choices as a goodness-of-fit measure.⁷⁰ The table reveals that with the dynamic pooled probit, of the 814 (= 693 + 121) observations for women who do not participate ($y = 0$), the dynamic pooled probit predicts 693 of these correctly (85.1 per cent). Similarly, the simple dynamic pooled probit predicts correctly 96 per cent [(= 1 – 107/2591)×100] for women who decide to participate ($y = 1$). The overall percentage that

⁶⁹ Specifically, in a footnote, McFadden (1977, p.35) states "...For example, values of 0.2 to 0.4 for R^2 represent excellent fit."

⁷⁰ See, for example, Wooldridge (2010, chapter 15.6, pages 573-574).

is correctly predicted is 93.5 per cent. The table for the standard dynamic random effects probit is similar, the overall percentage that is correctly predicted is 93.7 per cent. This leads us to the conclusion that the two models, the simple dynamic pooled probit and the

Table 5.2 The percent of choices correctly predicted

		Number of cases predicted			
		Pooled probit		RE probit	
		$y = 0$	$y = 1$	$y = 0$	$y = 1$
Number of actual cases	$y = 0$	693	121	681	133
	$y = 1$	107	2591	87	2611

standard dynamic random effects probit, do not differ substantially with respect to predictive accuracy.

Table 5.3 reports estimates of the coefficients for the dynamic models, of Heckman (1981b), Hyslop (1999), Keane and Sauer (2009), and Bartolucci and Nigro (2010). Column [1] gives the coefficient estimates of Heckman's (1981b) approach and the corresponding standard errors. His model has been estimated with the use of the GHK method and the variance-covariance matrix Σ of coefficients has been estimated by a sandwich formula.⁷¹ Heckman's (1981b) approach permits the presence of individual heterogeneity and the endogeneity of the initial conditions, but assumes no autocorrelation in the error term, ε_{it} .

As we have mentioned in section 3.3.1.1, in Heckman's (1981b) approach, exogeneity of the initial conditions can be considered as resulting from imposing the restriction $\theta = 0$ on equation (3.16). The estimate of θ is 1.12 and significantly greater than zero at the 5% level. This result strongly rejects the exogeneity of the initial conditions. The rejection implies that the time-invariant unobserved individual heterogeneity a_i is correlated with the initial period.

Compared with the standard dynamic random effects probit estimator, which treats the initial conditions as exogenous (column [2] in table 5.1), the Heckman (1981b) approach gives a slightly smaller coefficient estimate for state dependence (γ) of 1.98,

⁷¹ As Geyer (2013) states, "the asymptotic variance $\hat{f}_n(\hat{\theta}_n)^{-1} \hat{V}_n(\hat{\theta}_n) \hat{f}_n(\hat{\theta}_n)^{-1}$ is called the sandwich estimator, the metaphor being that $\hat{V}_n(\hat{\theta}_n)$ is a piece of ham between two pieces of bread $\hat{f}_n(\hat{\theta}_n)^{-1}$ ", in the case of model misspecification (p. 15).

compared to 2.02. Stewart (2007, p. 516) confirms this result when he compares the standard dynamic random effects probit model with Heckman's (1981b) approach by stating: "If the initial conditions are correlated with α_i ... this method of estimation overstates state dependence." Compared with the standard dynamic random effects probit estimates for the independent variables x_{it} , Heckman's (1981b) approach further increases the absolute values of the slope coefficients β .

At the end of column [1], the estimate of the standard deviation σ_α is also reported (referred to as Sigma in table 5.3). It indicates that 27.66% [= $0.6184^2 / (0.6184^2 + 1)$] of the unexplained variance can be attributed to the variance in the time invariant individual heterogeneity α_i . The Wald χ^2 statistic for the joint significance of the explanatory variables in the main equation (3.15) is 89.1 with 16 degrees of freedom (p -value < 0.001). Therefore, the null hypothesis that all coefficients, other than the constant, are jointly zero can be rejected even at the 1% level of significance. The value of the log-likelihood at convergence is -762.259.

Column [2] in Table 5.3 presents the Hyslop (1999) estimates. Hyslop (1999) introduces, apart from the individual heterogeneity and endogenous initial conditions, autoregressive error terms. The estimates are obtained with the use of the GHK method. Hyslop's (1999) approach, in column [2], produces coefficients on all the x_{it} variables that are smaller in absolute value than those of Heckman's (1981b) estimator, except for transitory income and whether the partner is unemployed.

Table 5.3 Dynamic probit models: Alternative approaches

	Heckman [1]	Hyslop [2]	Keane and Sauer [3]	Bartolucci and Nigro [4]
Constant	-0.7289 (0.9546)	-0.6641 (0.8023)	-0.7043 (0.8240)	
Participation lagged	1.9845*** (0.1583)	2.4003*** (0.1978)	2.3625*** (0.1857)	2.9402*** (0.2718)
Limiting Health Problems	-0.4051** (0.1580)	-0.3422** (0.1487)	-0.3484** (0.1491)	-0.3941 (0.2656)
Excellent Health	0.7340*** (0.2035)	0.7323*** (0.1990)	0.7328*** (0.1982)	0.5926 (0.4435)
Very Good/Good Health	0.5599*** (0.1674)	0.5549*** (0.1674)	0.5652*** (0.1666)	0.6632** (0.2985)
Fair Health	0.4784*** (0.1541)	0.4601*** (0.1570)	0.4781*** (0.1572)	0.8541*** (0.2950)
Arms, legs, hands, etc.	-0.0382 (0.1134)	-0.0205 (0.1030)	-0.0349 (0.1058)	-0.0846 (0.2809)
Anxiety/Depression	-0.3779** (0.1647)	-0.3481** (0.1524)	-0.3662** (0.1546)	-0.6979 (0.4500)
Years of Experience	0.0437*** (0.0084)	0.0320*** (0.0083)	0.0338*** (0.0084)	-0.0763 (0.0503)
Years of Education	0.0955*** (0.0308)	0.0759*** (0.0293)	0.0798*** (0.0293)	0.2230** (0.1059)
Presence of children aged 0-2 years	-0.7248** (0.2985)	-0.5747** (0.2823)	-0.6008** (0.2825)	-1.6976*** (0.5962)
Presence of children aged 3-4 years	-0.6197** (0.2441)	-0.4410* (0.2570)	-0.4713* (0.2576)	-1.1658*** (0.3998)
Presence of children aged 5-11 years	-0.3585*** (0.1060)	-0.2908*** (0.0977)	-0.2968*** (0.0979)	-0.0319 (0.2669)
Transitory income	-0.0433 (0.0690)	-0.0467 (0.0652)	-0.0456 (0.0655)	-0.0524 (0.1252)
Permanent income	-0.5218*** (0.1517)	-0.4575*** (0.1382)	-0.4709*** (0.1388)	
Spouse unemployed	-0.1598 (0.1442)	-0.1599 (0.1300)	-0.1519 (0.1280)	-0.5353** (0.2517)
Unemployment rate	0.0736*** (0.0240)	0.0484** (0.0230)	0.0540** (0.0230)	-0.0996* (0.0550)
Theta	1.1179** (0.4347)	1.3159* (0.7504)	0.8514* (0.4624)	
Sigma	0.6184*** (0.1227)	0.4574*** (0.1483)	0.4969*** (0.1333)	
ρ		-0.3164*** (0.0544)	-0.3417*** (0.0528)	
tau			-0.3369 (0.5286)	
lnL	-762.259	-751.045	-748.847	-254.680
BIC	1,837.4	1,814.9	1,810.5	720.2
AIC	1,600.5	1,578	1,573.6	569.3

Note: * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.
Standard errors are in parentheses.

Hyslop's (1999) approach gives a noticeably higher estimate of state dependence ($\hat{\gamma} = 2.40$), compared to Heckman's (1981b) approach ($\hat{\gamma} = 1.98$), with a considerably increased standard error. This result is in line with Michaud and Tatsiramos (2008, p. 17), who point out that "ignoring serial correlation leads to significantly lower state dependence effect" and with Stewart (2007, p. 525) when they compare a probit model with serial correlation in the error term accounted for to a model without serial correlation.

The estimate of θ is 1.31 and significantly greater than zero at the 10% level. This result rejects the exogeneity of the initial conditions again. We conclude that time-invariant unobserved individual heterogeneity α_i is correlated with the initial period.

The estimate of the standard deviation σ_α indicates that 17.28% of the total unexplained variance can be attributed to the variance of the time invariant individual heterogeneity α_i . Notice that the unexplained variance due to α_i gets smaller in this approach, compared with Heckman's (1981b) approach. The estimate of ρ in equation (3.18), labeled as ρ in the table, is significantly negative at the 1% level. The result implies that successive realizations of idiosyncratic errors ε_{it} are negatively correlated. The results imply that the effect of individual heterogeneity α_i is mitigated and decreases by around one third (compared to those in Heckman's 1981b approach) when we allow the idiosyncratic error ε_{it} to be serially correlated. This in turn implies that failure to control for serially correlated transitory errors would underestimate state dependence. The value of the log-likelihood at convergence is -751.045.

In column [3], we present our estimates based on Keane and Sauer's (2009) approach (see equation 3.20). As we described in chapter 3 (subsection 3.3.1.3), Keane and Sauer (2009) further generalized Hyslop (1999) and introduced a more flexible treatment of the initial conditions equation (i.e., the "reduced form equation") in which they defined an additional parameter, τ .

Compared with the estimates produced by Hyslop's (1999) approach, this approach slightly increases all coefficients in absolute value, except for the transitory non-labour income and whether the spouse is unemployed. State dependence, in this case, is 2.36. The estimate of θ gets smaller than in Heckman's (1981b) and in Hyslop's (1999) approaches and equals 0.85, but is significant at the 10% level. The standard deviation σ_α is 0.49 and implies that 19.8% of the total unexplained variance is attributed to the variance of α_i . The

estimate of ρ from equation (3.18) is smaller than in Hyslop's (1999) approach, but remains statistically significant at the 1% level. This indicates negative autocorrelation in ε_{it} in equation (3.18). The correlation coefficient τ , labeled as tau in the table, is -0.33 and not significant, suggesting that the Hyslop (1999) and the Keane and Sauer (2009) approaches do not differ significantly in this application. The value of the log-likelihood at convergence is -748.84 .

If we compare the effect of state dependence in these three approaches in table 5.3 (columns [1], [2] and [3]), there is a clear indication that the absence of autocorrelation in the error terms ε_{it} in Heckman's (1981b) approach underestimates the effect of state dependence. Moreover, we notice that Hyslop's (1999) model exhibits the largest estimate of state dependence and is larger than the Heckman (1981b) estimate by 15 per cent. In addition, overall, the Hyslop (1999) and Keane and Sauer (2009) models exhibit small differences in the estimates and, in most cases, both differ from those of the Heckman (1981b) model considerably.

Bartolucci and Nigro (2010) deal with the initial conditions problem via a fixed-effects approach and estimate the regression coefficients consistently without imposing distributional assumptions on the unobserved heterogeneity a_i . They eliminate a_i by using a suitable sufficient statistic. Nevertheless, a considerable pitfall of this approach for dynamic panel data analysis is that it requires that at least a transition between the states $y = 0$ and $y = 1$ takes place for an observation to contribute to the maximum likelihood estimation.⁷² Practically this means that conditioning on those observations which make a transition, $(0, 1)$ or $(1, 0)$, and discarding those which do not, $(0,0)$ or $(1,1)$, implies that model identification is based only on those observations where the dependent variable y_{it} changes over time. As a result, the number of usable observations is reduced drastically. In our analysis, valid observations are reduced from 3,865 (column [3] in Table 5.1) to 1,129 (from 331 to 108 women). Nevertheless, the Bartolucci and Nigro (2010) estimator can still be used to investigate female labour supply in a conditional fixed-effects approach. Column [4] provides the estimates of this approach which differ noticeably from those of columns [1]-[3]. The coefficient of the unemployment rate now becomes negative,

⁷² As was mentioned in section 3.3.2, Baltagi (2005, ch. 11, p. 210) demonstrates mathematically that when the dependent variable does not change over time observations add nothing to the likelihood.

thus supporting the “discouraged worker hypothesis.” All standard errors in column [4] are appreciably larger and in some estimates at least as twice as those in the previous three random effects approaches. The coefficient of excellent health has the expected (positive) sign, its magnitude decreases compared with the other three approaches, and becomes insignificant. Limiting health problems and problems related to anxiety and depression become insignificant, too. Finally, the high degree of state dependence is confirmed, as it increases considerably ($\hat{\gamma} = 2.94$), and stays significant at the 1% level. Note that the output does not provide an estimate for permanent non-labour income, as this variable has no variation over time. Wooldridge (2009), for linear models, states: “The fixed effects estimator allows for arbitrary correlation between α_i and the explanatory variables in any time period, just as with first differencing. Because of this, any explanatory variable that is constant over time for all i gets swept away by the fixed effects transformation” (page 481). Kohler and Kreuter (2008, p. 245) describe that “the fixed-effects model controls for all time-invariant differences between the individuals, so the estimated coefficients of the fixed-effects models cannot be biased because of omitted time-invariant characteristics...[like culture, religion, gender, race, etc]. One side effect of the features of fixed-effects models is that they cannot be used to investigate time-invariant causes of the dependent variables. ...Substantively, fixed-effects models are designed to study the causes of changes within a person [or entity]. A time-invariant characteristic cannot cause such a change, because it is constant for each person.” Similarly, Bartolucci (2009, p.18) states “An important drawback, common to all fixed-parameters approaches, is that the regression parameters for the time-constant covariates are not estimable.” The value of the log-likelihood at convergence is -254.680, which is much larger (algebraically) than in the previous three RE models, indicating a better fit.

In the last two lines of table 5.3, we include the values of the Bayes information criterion (BIC) and the Akaike information criterion (AIC). BIC and AIC are model selection methods and their value indicates which model improves forecasting accuracy (the smaller their value, the better the forecasting accuracy), provided that the datasets are identical. The BIC is defined as $BIC = -2\ln L + k\ln N$, where $\ln L$ is the maximized log-likelihood of the model, k is the number of parameters to be estimated and N is the number of observations. The AIC can be expressed as $AIC = -2\ln L + 2k$ (StataCorp., 2007). For the random effects approaches, the number of parameters to be estimated (the sum of

parameters from the two equations (3.15) and (3.16)) is 37 in [1], 38 in [2], and 39 in [3] and the valid number of observations is 3,767. For the fixed-effects approach, the number of parameters to be estimated is 29 and the valid number of observations is 1,129. In columns [2] and [3] of the table, the difference in BIC between Hyslop's (1999) approach and Keane and Sauer's (2009) approach is less than 10, indicating that the two approaches do not differ substantially with respect to predictive accuracy. The same applies to the difference in AIC criteria. In column [4], BIC and AIC criteria are estimated with a different size of dataset, therefore any comparison to the criteria from the previous approaches is not appropriate.

Among the three random effects approaches (Heckman [1], Hyslop [2] and Keane and Sauer [3]), the Keane and Sauer (2009) model appears to fit the data better than Heckman's (1981b) and Hyslop's (1999) models (see Table 5.4). The Keane and Sauer (2009) model produces a higher (algebraically) value of the log-likelihood function (-748.847) than the other two models. The additional parameter between [1] and [2] is the correlation coefficient ρ (see equation (3.18) in [2]) and the null hypothesis is that ρ is zero. Hence, the approximate χ^2 test has one degree of freedom. The difference in fit between models [1] and [2] is 22.42 and highly significant at the 1% level (p -value < 0.001). Similarly, the additional parameter between [2] and [3] is the coefficient τ (see equation (3.20) in [3]) and the null hypothesis is that τ is zero. The difference in fit between models [2] and [3] is 4.39 and statistically significant at the 5% level (p -value = 0.036). Although model [3] is significant, there is one reason that we tend to prefer model [2] rather than [3]: Models [2] and [3] produce, quantitatively, similar estimates, whereas Hyslop's (1999) model exhibits a greater effect of state dependence. Therefore, we are in favour of model [2]. These comparisons are shown in the following table.

Table 5.4. Comparisons of the results reported in columns [1], [2], and [2] of Table 5.3

Model		Log-likelihood	Models compared	Difference in Fit	
				-2(LL1-LL2)	p-value
Heckman	[1]	-762.259			
Hyslop	[2]	-751.045	[1] – [2]	22.428	<=0.001***
			[2] – [3]		
Keane and Sauer	[3]	-748.847		4.396	0.036**

Notes: ** significant at the 5 percent, *** significant at the 1 percent.

Table 5.5 presents the estimates of the initial conditions. As we have discussed in chapter 3, the reduced-form equation (see equation 3.16) accounts for the initial conditions, and is used to estimate the probability of participating in the first year of the sample. Here, the initial conditions equation includes the variable AGE as an exogenous time-varying instrumental variable, which does not appear in the main equation (3.15) and in the vector of the x_{it} independent variables. As we can see in the table, according to Heckman's (1981b), Hyslop's (1999) and Keane and Sauer's (2009) approaches, the variable AGE has a highly significant negative effect on the probability of participation at the 5% level. We also consider the variable "years of education" to be exogenous, because we assume that given that we start the analysis at age 39 for women, education has been completed by that age. Indeed, there are only minor changes in years of education after age 39 (see table 4.2).⁷³ The estimate of the years of education is also significant at the 5% level in all approaches. Limiting health problems are highly significant in Heckman's (1981b), Hyslop's (1999), Keane and Sauer's (2009) and Bartolucci and Nigro's (2010) approaches and years of experience at the 1% level, except for the Bartolucci and Nigro (2010) approach. Also, the presence of children in all age groups reduces the probability of participation, more noticeably when the children are aged 0-2 years old. Transitory income is statistically significant in all approaches, and whether the spouse is unemployed and the national unemployment rate are significant in the first three approaches.

⁷³ See, for example, Eckstein and Lifshitz (2011, footnote 14, p. 1688) for a similar approach, which considers education to be exogenous.

Table 5.5 Initial conditions equations

	Heckman [1]	Hyslop [2]	Keane and Sauer [3]	Bartolucci and Nigro [4]
Constant	-6.0106 (6.3424)	-6.4710 (6.2308)	-5.4332 (5.9336)	12.0681 (7.7819)
Limiting Health Problems	-1.7023*** (0.4853)	-1.6757*** (0.4773)	-1.6049*** (0.4354)	-2.1592** (1.0453)
Excellent Health	-0.1880 (0.7487)	-0.1616 (0.7556)	-0.1463 (0.6947)	-0.6289 (1.2381)
Very Good/Good Health	-0.3722 (0.7246)	-0.2922 (0.7311)	-0.3285 (0.6729)	-1.3624 (1.0489)
Fair Health	-0.0665 (0.7421)	-0.1480 (0.7503)	-0.1540 (0.6852)	-1.6552* (0.9083)
Arms, legs, hands, etc.	-0.5506 (0.3716)	-0.4569 (0.3735)	-0.4524 (0.3278)	0.2859 (0.5802)
Anxiety/Depression	-0.6289 (0.5823)	-0.5896 (0.5660)	-0.5966** (0.5262)	-0.7178 (0.8743)
Years of Experience	0.1224*** (0.0258)	0.1194*** (0.0262)	0.1077*** (0.0218)	-0.0585 (0.0531)
Years of Education	0.1444** (0.0645)	0.1512** (0.0626)	0.1231** (0.0549)	0.2489** (0.1100)
Presence of children aged 0-2 years	-6.1561*** (0.8234)	-9.2146*** (1.1053)	-8.8563*** (1.1205)	18.5955*** (1.3669)
Presence of children aged 3-4 years	-2.1913*** (0.4870)	-2.2374*** (0.4946)	-2.0438*** (0.4330)	19.9600*** (1.0665)
Presence of children aged 5-11 years	-1.1755*** (0.2950)	-1.1215*** (0.2912)	-1.0654*** (0.2559)	19.1016*** (0.6742)
Transitory income	-0.5820** (0.2776)	-0.5862** (0.2799)	-0.5190** (0.2527)	-0.7263* (0.3955)
Permanent income	-0.4466 (0.2963)	-0.4931* (0.2880)	-0.4142 (0.2650)	
Spouse unemployed	-0.8903* (0.5028)	-0.8949* (0.5101)	-0.8010* (0.4551)	0.1528 (0.2090)
Unemployment rate	1.4170** (0.5535)	1.4494*** (0.5381)	1.2927** (0.5044)	-1.7827 (1.4904)
Age	-0.1576** (0.0617)	-0.1497** (0.0631)	-0.1392** (0.0560)	

Notes: * denotes statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level. Standard errors are in parentheses.

Table 5.6 illustrates the average partial effects (APEs)⁷⁴ for the random effects probit that have been produced by using the approaches proposed by Heckman (1981b), Hyslop (1999), and Keane and Sauer (2009). The results obtained by Hyslop's (1999) approach are quite different in magnitude from those obtained by Heckman's (1981b) approach and similar to the results obtained by Keane and Sauer's (2009) approach. Standard errors are derived via the Delta method (see Wooldridge, 2010, Chapter 15 for an in-depth discussion; Lucchetti and Pignini, 2017, p. 17). For the fixed-effects approach of Bartolucci and Nigro (2010), as we have already mentioned in section 3.4, it is not technically feasible to calculate APEs (Lucchetti and Pignini, 2017, top of p. 16).

In all three approaches, the statistical significance of the vast majority of regressors is quite pronounced. Our primary interest is in the impact of health variables on the decision to participate. The APEs of the health-related variables are of the expected sign. In column [1], we present the APEs according to Heckman's approach. The presence of limiting health problems reduces the probability of participation by 0.056, compared to women who do not claim limiting health problems. The effect is highly statistically significant at the 1 % level. Having excellent health significantly increases the probability of participation by 0.0862 compared to women with poor or very poor health, while reporting very good or good health and fair health significantly increases the probability of participation by 0.0726 and 0.0557, respectively. Also the presence of anxiety and depression decrease the probability of participation by 0.0514. All these effects are highly statistically significant at the 5 % level.

In the case of Hyslop's (1999) approach, limiting health problems reduce the probability of participation by 0.0409 (significant at the 5% level), compared to women who do not claim limiting health problems. Having excellent health increases the probability by 0.0767 (significant at the 1% level), reporting very good or good health and fair health increases the probability of participation by 0.0639 and 0.0473, respectively (both significant at the 1% level), compared to women with poor or very poor health.

⁷⁴ See section 3.4 of the thesis for the description of the APEs.

Table 5.6 Dynamic probit models: Average Partial Effects (APEs)

	Heckman [1]	Hyslop [2]	Keane and Sauer [3]
Participation (lagged)	0.4579*** (0.0523)	0.6037*** (0.0741)	0.5851*** (0.0677)
Limiting Health Problems	-0.0560** (0.0241)	-0.0409** (0.0196)	-0.0424** (0.0200)
Excellent Health	0.0862*** (0.0224)	0.0767*** (0.0205)	0.0778*** (0.0206)
Very Good/Good Health	0.0726*** (0.0227)	0.0639*** (0.0208)	0.0660*** (0.0209)
Fair Health	0.0557*** (0.0168)	0.0473*** (0.0156)	0.0498*** (0.0157)
Arms, legs, hands, etc.	-0.0047 (0.0142)	-0.0022 (0.0112)	-0.0038 (0.0117)
Anxiety/Depression	-0.0514** (0.0242)	-0.0413** (0.0199)	-0.0444** (0.0206)
Years of Experience	0.0054*** (0.0010)	0.0034*** (0.0009)	0.0037*** (0.0009)
Years of Education	0.0118*** (0.0037)	0.0008*** (0.0031)	0.0088*** (0.0032)
Presence of children aged 0-2 years	-0.1085** (0.0526)	-0.0738* (0.0424)	-0.0788* (0.0435)
Presence of children aged 3-4 years	-0.0903** (0.0415)	-0.0543 (0.0362)	-0.0594 (0.0374)
Presence of children aged 5-11 years	-0.0482*** (0.0150)	-0.0339*** (0.0120)	-0.0352*** (0.0122)
Transitory income	-0.0053 (0.0085)	-0.0050 (0.0070)	-0.0050 (0.0072)
Permanent income	-0.0648*** (0.0183)	-0.0496*** (0.0148)	-0.0519*** (0.0151)
Spouse unemployed	-0.0198 (0.0180)	-0.0173 (0.0142)	-0.0167 (0.0142)
Unemployment rate	0.0091*** (0.0029)	0.0052** (0.0025)	0.0059** (0.0025)

Notes: * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. Standard errors are in parentheses.

Reporting anxiety and depression diminishes the probability of participating by 0.0413 (significant at the 5% level). Problems related to arms, legs, hands, etc., do not reduce the probability of participation significantly, however.

As for the presence of children, the related variables have a significantly negative effect on the probability of women's participation in Heckman's (1981b) model, with the youngest children aged 0-2 years old having the strongest effect, namely -0.1085 (significant at the 5% level), compared to a woman who has no children in this age group. A child aged 3-4 reduces this probability by 0.0903 (significant at the 5% level), whereas a child who attends school reduces the probability by only 0.0482 (significant at the 1% level). The corresponding estimated partial effects in the Hyslop (1999) and Keane and Sauer (2009) approaches indicate that ignoring autocorrelation overestimates the coefficient estimates for the children related variables in Heckman's (1981b) approach. As we can see in columns [2] and [3], an additional child aged 0-2 reduces the probability of women's participation by 0.0738 and 0.0788 respectively, a child aged 3-4 by 0.0543 and 0.0594, respectively, whereas a child aged 5-11 by 0.0339 and 0.0352, respectively. Note that, assuming two-sided alternatives, which is the standard strategy and which we follow, the estimates obtained from the Hyslop (1999) and Keane and Sauer (2009) approaches for the coefficients of children in the age group 3-4 years old are insignificant even at the 10% level, as the corresponding t -statistics are $-0.0543/0.0362 = -1.50$ and $-0.0594/0.0374 = -1.59$.⁷⁵

Years of experience and years of education exhibit the expected sign in all approaches. Heckman's (1981b) approach gives the largest coefficient estimates for these two elements of human capital. According to the estimates from Heckman's (1981b) approach, an additional year of work experience significantly augments the probability of participation by 0.0054 and an additional year of education significantly augments the probability of participation by 0.0118 (both statistically significant at the 1% level).

⁷⁵ It would be more reasonable, however, to assume *one*-sided alternatives whenever the sign of a coefficient is predicted by economic theory, in which case these two coefficients would be statistically significant at the 10% level. By the same token, the coefficients for children in the age group 0-2 years old in these two models are significant at the 5% (and not at the 10%) level, as their t -statistics are $-0.0738/0.0424 = -1.74$ and $-0.0788/0.0435 = -1.81$.

Hyslop's (1999) and Keane and Sauer's (2009) approaches give smaller APEs, but still significant at the 1% level.

The estimated coefficients on monthly non-labour income, whether it is transitory or permanent, have the expected negative sign. In Heckman's (1981b) approach, the APEs indicate that the probability of participation decreases by 0.0648 as the permanent non-labour income increases. According to Hyslop's (1999) model, this probability declines by 0.0496, and according to Keane and Sauer's (2009) model, it declines by 0.0519. Note, however, that although permanent non-labour income is statistically significant (at the 1% level) in all three approaches, transitory non-labour income is not.

Finally, surprisingly and opposite to Heckman (1981a), we find a positive impact of the national unemployment rate on participation decisions. The unemployment rate raises the probability of participation in all three approaches significantly. A possible explanation might be that, in theory, the total unemployment rate (i.e., men and women) might also have a positive effect on the labour force participation of women, other things being equal. For example, whenever the national unemployment rate is high, more husbands are unemployed, and this might motivate more wives, who would not otherwise work, to enter the labour market (Hatzinikolaou, 2018, p.38).

In all three approaches, the coefficient estimate on lagged participation reflects a high degree of positive state dependence. The estimates of the APEs, 0.4579, 0.6037, and 0.5851 (all significant at the 1% level) in the first row of the table 5.6, clearly indicate that there is a positive and highly significant correlation between participation at current year and participation at the previous year. Notice that the estimate obtained from Hyslop's (1999) approach dominates the other two in magnitude. It shows that, averaged across all women and all time periods, the probability of a woman participating in the current year is higher by 0.6037 if the woman was participating in the previous year, compared to a woman who was not participating in the previous year.

The following tables (table 5.7 and table 5.8) report predicted probabilities of participating and estimates of the state dependence for the year 2003.⁷⁶ The third and fourth columns report the average estimated probabilities of participating in 2003 given that the

⁷⁶ We implement the coding proposed by Lucchetti and Pigni (2017), who follow the methodology of Wooldridge (2005).

woman was or was not a participant in 2002, for the presence of children aged under 11

Table 5.7 Estimated probabilities of participating (APEs) for the presence of health problems, the presence of children, and state dependence for 2003, according to Heckman's (1981b) model

Estimated probabilities according to Heckman's (1981b) model				
Presence of children	Presence of health problems	y = 1 in 2002	y = 0 in 2002	State dependence
Children < 11 years old	without	93.00%	48.30%	44.70%
Children < 11 years old	with	61.40%	10.50%	50.90%
Children <= 2 years old	without	82.30%	27.20%	55.10%
Children <= 2 years old	with	39.30%	3.40%	35.90%

Table 5.8 Estimated probabilities of participating (APEs) for the presence of health problems, the presence of children, and state dependence for 2003 according to Hyslop's (1999) model

Estimated probabilities according to Hyslop's (1999) model				
Presence of children	Presence of health problems	y = 1 in 2002	y = 0 in 2002	State dependence
Children < 11 years old	without	95.20%	35.60%	59.60%
Children < 11 years old	with	67.30%	5.50%	61.80%
Children <= 2 years old	without	88.10%	19.60%	68.50%
Children <= 2 years old	with	48.30%	1.80%	46.50%

years old, the presence of children aged up to 2 years old and the presence of health problems. In this part of the analysis, we have summed up as 'health problems' the estimates of the main significant health problems, that is, poor or very poor health, health problems that limit daily activities, and health problems related to anxiety and depression. Hence, health problems act cumulatively.

Table 5.7 reports the predicted probabilities and estimates of state dependence produced by Heckman's (1981b) model. For a married woman who participated in 2002, with children under 11 years old, and without health problems, the estimated probability of participating in 2003 is 93 percent. For a married woman with the same characteristics who did not participate in 2002, the estimated probability of participating in 2003 is 48.30 percent. The difference in estimated probabilities, that is, 44.70 percentage points, is an estimate of the state dependence of participating. The presence of health problems for the same woman decreases the probabilities of participating in 2003 by 61.40 percent and 10.50 percent, respectively. With presence of children less than 2 years old and health problems, the estimated probabilities are markedly diminished. The last line of table 5.7 clearly depicts this decrease. For a married woman who participated in 2002, with children aged below 2 and health problems, the estimated probability of participating in 2003 is now 39.3 percent. If she did not participate in 2002, the estimated probability of participating in 2003 becomes only 3.4 percent. Consequently, the estimate of the state dependence falls to 35.9 percent.

Columns 3 and 4 of table 5.8 report the corresponding estimated probabilities of participating in 2003 given that the woman was or was a participant in 2002 according to Hyslop's (1999) model. The last column contains estimates of the magnitude of the state dependence for the year 2003. The results are similar to those from Heckman's (1981b) model, except for the more diminishing effect of health problems on the predicted probabilities and, consequently, on the estimates of state dependence, in the case a woman did not participate in 2002. Specifically, the probability of a woman, with young children aged below 2 and health problems, participating in 2003 falls to 1.80 percent, compared to a woman who did not participate in 2002. State dependence estimates in table 5.8 (last column) are very close to those reported in table 5.6 according to Hyslop's (1999) approach (where $\hat{\gamma} = 0.6037$). Overall, both tables 5.7 and 5.8 demonstrate separately the strong effect of the presence of young children and of health problems on the probability of participating.

Table 5.9 presents the results produced by the Mundlak (1978) specification, embedded in Hyslop's (1999) approach. According to Mundlak (1978), we assume a linear relationship between the unobserved heterogeneity α_i and the means of all time-

Table 5.9 Hyslop's approach with correlated random effects (Mundlak's approach)

	Hyslop with Mundlak [1]	Average partial effects (APEs) [2]	Initial Conditions [3]
Constant	-0.2517 (0.8113)		-5.6396 (5.8588)
Participation lagged	2.3801*** (0.1832)	0.5996*** 0.0702	
Limiting Health Problems	-0.2263 (0.1633)	-0.0265 (0.0205)	-1.2332** (0.4997)
Limiting Health Problems (mean)	-1.0827*** 0.3413	-0.1194*** (0.0378)	-1.3554** (0.5613)
Anxiety/Depression	-0.3506** (0.1519)	-0.0424** (0.0202)	-0.4010 (0.5307)
Year of Experience	0.0332*** 0.0085	0.0036*** (0.0009)	0.1174*** 0.0238
Years of Education	0.0763*** (0.0288)	0.0084*** (0.0031)	0.1439** (0.0602)
Presence of children aged 0-2 years	-0.6264** (0.2808)	-0.0834* (0.0444)	-7.6678*** (0.6480)
Presence of children aged 3-4 years	-0.4539* (0.2561)	-0.0571 (0.0370)	-2.1936*** (0.4533)
Presence of children aged 5-11 years	-0.2918*** (0.0977)	-0.0347*** (0.0123)	-1.0104*** (0.2629)
Transitory income	-0.0700 (0.0641)	-0.0077 (0.0070)	-0.7615*** (0.2534)
Permanent income	-0.3435*** (0.1265)	-0.0379*** (0.0138)	-0.4030 (0.2760)
Spouse unemployed	-0.3823** (0.1841)	-0.0421** (0.0203)	-1.3400*** (0.4022)
Unemployment rate	0.0481** (0.0229)	0.0053** (0.0025)	1.3272*** (0.5057)
Age			-0.1486** 0.0578
Theta	1.0613** (0.5217)		
Sigma	0.4767*** (0.1377)		
ρ	-0.3107*** (0.0527)		
lnL	-760.452		
BIC	1,727.9		
AIC	1,580.9		

Note: * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. Standard errors are in parentheses.

varying explanatory variables $\mathbf{x}_{it}$ (see chapter 3, section 3.3.3). Our purpose is to check whether health-related variables, self-assessed health status, health limitations and anxiety/depression, are associated with unobserved heterogeneity that reduces or increases the probability of participating.⁷⁷

We began our analysis with all health-related variables included among the regressors, but the results demonstrated that when self-assessed health and limiting health problems coexist among the regressors, the effect of self-assessed health is no longer significant, due to the high correlation between self-assessed health and limiting health problems. That is, the variance shared by self-assessed health and participation is due solely to the variability of the variable “limiting health problems.” Consequently, the three dummy variables for self-assessed health are removed from the model. Further, we removed the highly insignificant variables, one at a time, re-estimated the model and kept those variables in the regression that have the correct sign and have a z -statistic larger than or equal to 1 (in absolute value).

Table 5.9 reports estimates of the coefficients of the model. The estimates are obtained with the use of the GHK method and the variance-covariance matrix Σ of coefficients has been estimated by a sandwich formula. In column [1], the estimate of state dependence γ is 2.38, is very close to that from Hyslop’s (1999) approach ($\hat{\gamma} = 2.40$ in table 5.3), and is highly significant at the 1% level. The estimated coefficient on the time mean of health limitations, labeled as “Limiting Health Problems (mean),” is significantly different from zero at the 1% level, which suggests that the exogeneity of the health-related variable is rejected. The estimated coefficient is -1.0827 and indicates that limiting health problems are correlated with individual characteristics that reduce the probability of participation.

Column [2] in table 5.9 illustrates the APEs. The presence of limiting health problems (the mean) reduces the probability of participation by 0.1194, compared to women who do not claim limiting health problems. The effect is statistically significant at the 1% level. Also the presence of anxiety and depression decreases the probability of participation by 0.0424 at the 5% level. Years of experience and years of education exhibit

⁷⁷ We thank Claudia Pigni (Department of Management, Polytechnic University of Marche, Italy) for correspondence upon the topic.

the expected positive sign (significant at the 1% level) and the estimated coefficient on monthly permanent non-labour income has the expected negative sign (significant at the 1% level). The APEs indicate that the probability of participation decreases by 0.0379 as the permanent non-labour income increases. In addition, we find a negative impact of whether the spouse is unemployed (significant at the 5% level). Its sign signifies a negative added worker effect and implies that a woman whose partner is unemployed is more likely to be non-participant than a woman whose partner is employed. Our finding verifies empirical studies, which find that the added worker effect is negative in the UK. Moreover, we find a positive impact of the national unemployment rate on participation decisions (significant at the 5% level). In addition, this approximation supports the assumption that children affect women's labour supply. The relevant estimates of the coefficients for the presence of children aged 0-2 and aged 5-11 years old have the expected negative sign, and are statistically significant.

Finally, in column [2], the coefficient estimate on lagged participation reflects a high degree of positive state dependence. The estimate of the APEs is 0.5996 (significant at the 1% level) in the first row of the table 5.9, and clearly indicates that there is a positive and highly significant correlation between participation at current year and participation at the previous year. The estimate is almost identical to the estimate obtained from Hyslop's (1999) approach without Mundlak's specification ($\hat{\gamma} = 0.6037$ in table 5.6). Finally, column [3] presents the estimates for the initial conditions equation (see equation 3.16).

Table 5.9 also includes the standard deviation of the individual heterogeneity σ_α (denoted as Sigma) and is equal to 0.4767. Subsequently, the estimated value of the intra-group correlation coefficient equals 0.3884 $[= 0.4767/(0.4767^2 + 1)]$ and implies that 38.84 per cent of the total unexplained variation in participation can be attributed to the unobserved individual heterogeneity ξ_i (see equation (3.36)). The estimate of ρ (see equation (3.18)) equals -0.3107 and is significantly negative at the 1% level. The result implies that successive realizations of idiosyncratic errors ε_{it} are negatively correlated. Finally, the estimate of θ is 1.06 and is significantly greater than zero at the 5% level. This result rejects the exogeneity of the initial conditions and indicates that time-invariant unobserved individual heterogeneity is correlated with the initial period.

The BIC and AIC criteria have values 1,727.9 and 1,580.9, respectively. Compared with the corresponding values obtained from Hyslop's (1999) approach (in table 5.3, the

values are 1,814.9 and 1,578 for the BIC and AIC criteria, respectively), we notice that the value of the BIC criterion in this approach is less than the value of the BIC criterion in Hyslop's (1999) approach (the difference in values of BIC criterion is approximately 79), whereas the difference in values of AIC criterion is approximately 3 between the two approaches. This indicates that the Mundlak (1978) specification, embedded in Hyslop's (1999) approach, appears to fit the data better than the approach of Hyslop (1999), suggesting that individual characteristics are correlated with limiting health problems.

5.3 Summary

The analysis demonstrated that state dependence is a statistically significant ingredient in women's labour force participation decision in all approaches. Among the Heckman (1981b), Hyslop (1999) and Keane and Sauer (2009) approaches, Hyslop's (1999) approach provides the largest estimate of the coefficient on state dependence ($\hat{\gamma} = 2.40$). Consequently, Hyslop's (1999) approach significantly augments the average partial effect for state dependence, which suggests that the probability of women's participation increases by 60.37 per cent at current year for women who were participating at the previous year, compared to women who were not in the labour market at the previous year (see Table 5.6). In addition, Hyslop's (1999) approach provides us with clear evidence that there is autocorrelation in the unobserved time-varying idiosyncratic error term ε_{it} . Moreover, Hyslop's (1999) model shows a significant higher fit than Heckman's (1981b) model, and unimportant loss of fit comparing with Keane and Sauer's (2009) model. Hence, we are in favour of Hyslop's (1999) approach.

The results also clearly show that health is a statistically significant determinant of women's labour force participation decisions. Women have a greater possibility of participation when they do not report limiting health problems and when they claim excellent, very good, good or fair health, compared to women who claim poor or very poor health. For all health-related variables, with the exception of health problems related to arms, legs, hands, etc. we observe a large and highly significant effect. In Heckman's (1981b), Hyslop's (1999) and Keane and Sauer's (2009) approaches, the health-related variables exhibit the expected sign and significance. In the fixed-effects approach of

Bartolucci and Nigro (2010), however, excellent health and presence of anxiety and depression become insignificant. Finally, Mundlak's (1978) approximation, embedded in Hyslop's (1978) approach, indicated that limiting health problems are correlated with unobserved characteristics that reduce the probability of participation.

The variables related to children are of the expected sign and statistically significant in all approaches. The presence of at least one young child in the household diminishes the probability of a woman's participation. The younger the children, the higher the effect (7.38, 5.43 and 3.39 percentage points for children aged 0-2, 3-4 and 5-11 years old, respectively, accordingly to Hyslop's (1999) approach). An additional year of work experience and an additional year of education increase noticeably the probability of participation. In addition, in all approaches, coefficient estimates for permanent non-labour income demonstrate statistical significance. As in Hyslop (1999), the transitory non-labour income does not appear to be a significant predictor of a woman's decision for labour force participation. Finally, the effect of the national unemployment rate on participation is found to be positive at the 1% level, according to Hyslop's (1999) approach.

Chapter 6

Conclusion

We attempted to explore the determinants of women's labour force participation decisions and primarily to investigate whether and to what extent health affects labour force participation decisions. Based on Hyslop's (1999) dynamic utility maximization model of female labour supply, i.e., the decision to participate in the labor market emerges from the solution of a dynamic programming problem and depends on the difference between the reservation wage and the market wage, we constructed a reduced form specification of labour supply, and we substituted the reservation wage by a function of individual characteristics and market conditions, and we also substituted the market wage by a function of market conditions that determine the market wage. In line with the above, we estimated dynamic discrete choice models within a framework determined by Heckman (1981b), Hyslop (1999), Keane and Sauer (2009), Bartolucci and Nigro (2010), and Mundlak (1978). In all approaches, we considered the participation status in the previous year as well health-related variables as main determinants in women's labour force participation decision at current year and accounted for the initial conditions and the presence of individual heterogeneity.

We used data from the first twelve waves of the British Household Panel Survey (BHPS) in the UK and, in line with Eckstein and Wolpin (1989), we constructed a balanced panel of 331 women aged between 39 and 45 years who are either continuously legally married or cohabitating across all the twelve years of the survey. The time span of the sample permitted us to investigate the determinants of labour supply, before women decide whether or not to retire. This special characteristic of our data distinguishes our analysis from the majority of studies that emphasize either decisions to retire or work decisions of older men in relation to their health. To the best of our knowledge, our two contributions to the literature are as follows: (1) it is the first study that examines the impact of both physical and mental health on women's labour force participation decisions as we use several aspects of health as proxies for health (in particular, self-assessed health, whether health limits daily activities and health problems related to anxiety and depression), and

(2) it is the first study that examines female labour supply within a framework determined by Bartolucci and Nigro (2010).

In all approaches, the coefficient of lagged participation reflects a high degree of positive state dependence and is statistically significant at the 1% level. The significance of individual heterogeneity in the random effects probit model indicates that individual heterogeneity is present, and the Heckman (1981 β) approach further implies that the individual heterogeneity is also correlated with the initial period. Comparing the effect of state dependence in the Heckman (1981b), Hyslop (1999), and Keane and Sauer (2009) approaches, we found a clear indication that the absence of autocorrelation in the error term in Heckman's (1981b) approach underestimates the effect of state dependence. As in Hyslop (1999), we found a significant negative autocorrelation in the error term.

The coefficients on health-related variables have the expected sign. Women have a greater probability of participation when they do not report limiting health problems and when they claim excellent, very good, good or fair health, compared to women who claim poor or very poor health. Mundlak's (1978) correlated random effects approach, embedded in Hyslop's (1999) approach, also showed that the estimate of state dependence is highly significant at the 1% level and close to that obtained from the Hyslop (1999) estimator. The exogeneity of the health-related variable "Limiting Health Problems (mean)" is rejected. The estimated coefficient indicates that limiting health problems are correlated with unobserved characteristics that reduce the probability of participation. Hence, our empirical analysis strengthens the view that health is an important determinant of women's labour force participation decisions.

The relevant literature also acknowledges that health is a multi-dimensional concept and its definition depends in part on the questions the researchers address. Furthermore, there is no consensus about the measurements of health that are more accurate as determinants of female labour supply. Consequently, both the definition of health and the measurements used in the empirical literature vary from study to study. Our research contributes to the literature, as it strengthens the understanding of the relationship between health and labour force participation decisions by incorporating four health-related variables: (i) self-assessed health, (ii) whether health limits daily activities, (iii) health problems related to arms, legs, hands, etc., and (iv) health problems related to anxiety and depression. Similarly to Jones et al. (2013), we consider that health limitations

to daily activities, in particular, serve as a more specific measure of disability and therefore may capture long-standing health problems that strongly influence individuals' work activities, whereas chronic illnesses captured by self-assessed health may not substantially limit work activities. We hope that the choice of these variables will shed additional light on the dynamic interaction between labour force participation and health for women.

In every model we estimate, the coefficients of all the other important variables are also of the expected sign. The presence of at least one young child in the household diminishes the probability of a woman's participation, and the younger the child the higher the effect. An additional year of work experience and an additional year of education increase significantly the probability of participation. Coefficient estimates for permanent non-labour income are also statistically significant. In contrast, transitory non-labour income does not appear to be a significant predictor of the decision to participate in the labour market, as we expected, since transitory changes in family income are not expected to cause changes in work habits. The added worker effect is found to be significant only in the Bartolucci and Nigro (2010) approach and Mundlak's (1978) approximation. Its negative sign implies that a woman whose partner is unemployed is more likely to be non-participant than a woman whose partner is employed. Our finding verifies empirical studies, which find that the added worker effect is negative in the UK. According to the suggestions of the examination committee, the negative added worker effect is not necessarily due to the complementarity of partners' leisure times, but is likely due to high unemployment benefits in the UK. Besides, the effect of the national unemployment rate on participation is found to be positive and significant in all models. One exception to the positive unemployment rate is found in the Bartolucci and Nigro (2010) approach. Finally, the predicted probabilities demonstrated the strong effect of the presence of young children and of health problems on the probability of participating. Specifically, the probability of a woman, with young children aged below 2 and health problems, participating in 2003 falls to 1.80 percent, compared to a woman who did not participate in 2002.

Our analysis suggests possible lines for future research. It would be of great interest to specify a dynamic model that examines the endogeneity of health by exploring whether dual causality exists between health and participation, i.e. whether health affects participation status and vice versa and whether there are unobserved individual effects that may affect health and work outcomes simultaneously. In the area of labour economics, the

vast majority of studies examine only the impact of health on participation status. The justification is that poor health is the principal risk that individuals encounter and participation status is the decision taken conditional on health status. On the other hand, in the area of health economics, the focus is on the impact of participation status on health, and participation status is among the other determinants that affect health; see, for example, Contoyannis, et al. (2004). Only a few studies take the dual causality into account and explore the endogeneity issue of health and participation (Haan and Myck, 2009, Lindeboom and Kerkhofs, 2009). This is a great motivation for a similar future research focused on women.

From a policy prospective, identifying the determinants of labour force participation decisions of women and, in principle, the barriers that women are likely to face over the life course, which are hidden behind the labour force participation decisions, can improve our understanding of the dynamics of female labour supply. At the same time, an ageing society imposes economic challenges for the economy. In order to address the economic challenges of an ageing society, the increase in employment rates are considered fundamental. For middle-aged women, there already exist policies aimed to engage women in paid employment and encourage those who do not work to enter the labour market. But for women who have health problems and for those who have experienced a health shock such policies are limited. There could be policies that promote uninterrupted employment trajectories for women, despite their health problems, including greater workplace flexibility to respond to their need for medical care, stress-free work environments for ill women and also, equally important, public actions targeted to prevent disease. Besides, a greater childcare provision for these women should be of a prime concern. The effectiveness of policies designed to support simultaneously employment and health of women depends on a better understanding of the impact of health on labour supply. The policies could also be applied for other disadvantaged and vulnerable population groups.

References

- Alessie, R., Hochguertel, S. and van Soest, A. (2004). Ownership of stocks and mutual funds: a panel data analysis. *Review of Economics and Statistics*, 86(3), 783–796.
- Altuğ, S., & Miller, R. (1998). The Effect of Work Experience on Female Wages and Labour Supply. *The Review of Economic Studies*, 65(1), 45-85. Retrieved from <http://www.jstor.org/stable/2567003>.
- Andriopoulou E. and Tsakloglou P. (2011). The Determinants of Poverty Transitions in Europe and the Role of Duration Dependence. *IZA Discussion Paper* 5692.
- Arber, S., Gilbert, G.N. and Dale, A. (1985). Paid employment and women's health: a benefit or a source of role strain? *Sociology of Health & Illness*, 7(3), 375-400.
- Arulampalam, W. and Stewart, M. B. (2009). Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators. *Oxford Bulletin of Economics and Statistics*, 71(5), 659–681.
- Arulampalam, W., Booth, A., & Taylor, M. (2000). Unemployment Persistence. *Oxford Economic Papers*, 52(1), 24-50.
- Baltagi, B.H. (2005). *Econometric Analysis of Panel Data* (3rd ed.). John Wiley & Sons Inc., New York.
- Bartolucci, F. (2019). *Seminar notes: Analysis of binary panel data by static and dynamic logit models*, Università di Napoli – Federico II, Italy.
Retrieved from: http://www.stat.unipg.it/~bart/index_file/Page473.html
- Bartolucci, F. and Nigro, V. (2010). A Dynamic Model for Binary Panel Data with Unobserved Heterogeneity Admitting a $\sqrt{n}$ -Consistent Conditional Estimator. *Econometrica*, 78(2), 719–733.
- Bartolucci, F. and Pignini, C. (2017). cquad: An R and Stata Package for Conditional Maximum Likelihood Estimation of Dynamic Binary Panel Data Models. *Journal of Statistical Software*, 78(7), 1 - 26. doi:<http://10.18637/jss.v078.i07>
- Becker, G.S. (1967). Human capital and the personal distribution of income, University of Michigan Press, Ann Arbor. Reprinted in G.S. Becker (1993) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (3rd ed.). University of Chicago Press (for NBER), Chicago.
- Μπερτσεκάς, Δ. Π. και Τσιτσικλής, Γ. Ν. (2016). Εισαγωγή στις Πιθανότητες με Στοιχεία Στατιστικής [Introduction to Probabilities with Statistics]. Μετάφραση: Αικατερίνη Χούστη, Εκδ. Τζιόλα, Θεσσαλονίκη.
- Biewen, M. (2009). Measuring State Dependence in Individual Poverty Histories When There Is Feedback to Employment Status and Household Composition. *Journal of Applied Econometrics*, 24(7), 1095-1116.

- Bound, J. (1991). Self-Reported Versus Objective Measures of Health in Retirement Models. *The Journal of Human Resources*, 26(1), 106-138. doi:10.2307/145718
- Breunig, I.M. (2011). *Three essays on health and labor outcomes (PhD thesis)*. Colorado State University, Colorado, USA. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1785633
- Brown, S., J. Roberts and K. Taylor (2010). Reservation Wages, Labour Market Participation and Health. *Journal of the Royal Statistical Society, Series A*, 173(3), 501-529.
- Cai, L. (2010). The Relationship between Health and Labor Force Participation: Evidence from a Panel Data Simultaneous Equation Model. *Labor Economics*, 17(1), 77-90.
- Cai, L. (2018). Dynamic Labour Supply of Married Australian Women. *LABOUR*, 32(3), 427-450.
- Cai, L. (2021). The effects of health on the extensive and intensive margins of labour supply. *Journal of the Royal Statistical Society*, 184(1), 87-117. doi: <https://doi.org/10.1111/rssa.12615>
- Cai, L., & Kalb, G. (2006). Health status and labour force participation: evidence from Australia. *Health economics*, 15 (3), 241–261.
- Campolieti, M. (2002). Disability and the labour force participation of older men in Canada. *Journal of Labor Economics*, 9(3), 405-432.
- Chay, K.Y. and Hyslop, D. (2001). *Identification and estimation of dynamic binary response models: Empirical evidence using alternative approaches* (Working Paper No. 5). Berkeley: Berkeley Center for Labor Economics, University of California.
- Chirikos, T. N. (1993). The Relationship Between Health and Labor Market Status. *Annual Review of Public Health*, 14(1), 293-312.
- Chirikos, T. N. and Nestel, G. (1981). Impairment and Labor Market Outcomes: A Cross-Sectional and Longitudinal Analysis. In Herbert S. Parnes (eds.), *Work and Retirement*. The MIT Press: Cambridge, MA.
- Clark, A., and Oswald, A. (1994). Unhappiness and Unemployment. *Economic Journal*, 104(424), 648–659.
- Collet, R. and Legros, D. (2016). Dynamics of female labour force participation in France. *Applied Economics*, 48(30), 2807-2821. doi: 10.1080/00036846.2015.1130787
- Contoyannis, P. and Rice, N. (2001). The impact of health on wages: Evidence from the British Household Panel Survey. *Empirical Economics*, 26, 599–622.
- Contoyannis, P., Jones, A.M. and Rice, N. (2004). The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics*, 19(4), 473-503.
- Cottrell, A. and Lucchetti, R. *GRET, User's Guide, Gnu Regression, Econometrics and Time-series Library*. Retrieved from <http://gretl.sourceforge.net/>

- Cox, D.R. (1972). The Analysis of Multivariate Binary Data. *Applied Statistics*, 21(2), 113–120.
- Currie, J. and Madrian, B.C. (1999). Health, health insurance and the labor market. In Ashenfelter, O. and Card, D. (Eds.), *Handbook of Labor Economics* (Volume 3, part C, pp. 3309-3416). Elsevier, Amsterdam.
- Del Boca, D. and R. M. Sauer (2009). Life cycle employment and fertility across institutional environments. *European Economic Review*, 53(3), 274–292.
- Del Boca, D., Locatelli, M. and Pasqua, S. (2000). Employment Decisions of Married Women: Evidence and Explanations. *LABOUR*, 14(1), 35:52.
- DeGroot, M. H. (1986). *Probability and Statistics* (2nd ed.). Addison-Wesley, Reading, Massachusetts.
- Delattre, E. and Moussa, R. (2018). On the Estimation of Causality in a Bivariate Dynamic Probit Model on Panel Data with Stata Software: A Technical Review. *Theoretical Economics Letters*, 8(6), 1257-1278. doi: 10.4236/tel.2018.86083.
- Delattre, E., Moussa, R. and Sabatier, M. (2019). Health condition and job status interactions: econometric evidence of causality from a French longitudinal survey. *Health Economics Review*, 9(3). doi: <https://doi.org/10.1186/s13561-019-0220-3>
- Dex, S. (1987). *Women's Occupational Mobility*, London: Macmillan.
- Doiron, D., Fiebig D. G., Johar, M. and Suziedelyte, A. (2015). Does self-assessed health measure health? *Applied Economics*, 47(2), 180-194.
- Eckstein, Z. and Wolpin, K. (1989). Dynamic labor force participation of married women and endogenous work experience. *The Review of Economic Studies*, 56(3), 375–390.
- Eckstein, Z., & Lifshitz, O. (2011). DYNAMIC FEMALE LABOR SUPPLY. *Econometrica*, 79 (6), 1675-1726.
- Edon, C. and Kamionka, Th. (2010). *Dynamic modeling of fertility and labour market participation of married or cohabiting women*. LIDAM Discussion Papers IRES, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).
- Eurostat (2020). Databrowser. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/HLTH_SILC_04_custom_281716/default/table?lang=en
- Faggian, A. (2014). Job search theory. In *Handbook of Regional Science* (pp. 59-73). Springer, Berlin, Heidelberg. https://www.researchgate.net/profile/Alessandra_Faggian/publication/278661248_Job_Search_Theory/links/570ce23c08aec783ddcd9f1e/Job-Search-Theory.pdf
- Francesconi, M. (2002). A Joint Dynamic Model of Fertility and Work of Married Women. *Journal of Labor Economics*, 20(2), 336-380. doi:10.1086/338220

- Friedman, M. (1957). *The Permanent Income Hypothesis. A Theory of the Consumption Function*. Princeton University Press. ISBN 978-0-691-04182-7.
- García-Gómez, P., Jones, A. M. and Rice, N. (2010). Health effects on labour market exits and entries. *Labour Economics*, 17(1), 62–76.
- Garibaldi, P. and Wasner, E. (2003). *Equilibrium Employment in a Model of Imperfect Labor Markets*. IZA Discussion Paper No. 950.
- Ge, S. (2011). Women's College Decisions: How Much Does Marriage Matter? *Journal of Labor Economics*, 29 (4), 773-813.
- Geyer, C. J. (2013). *5601 Notes: The Sandwich Estimator*. University of Minnesota, School of Statistics, USA.
- Geweke, J. (1988). Antithetic acceleration of Monte Carlo integration in Bayesian inference. *Journal of Econometrics*, 38(1-2), 73-89.
- Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2), 223-255.
- Haan, P. and Myck, M. (2009). *Dynamics of Poor Health and Non-employment*. IZA Discussion Paper No. 4154.
- Hajivassiliou, V. (1993). Simulation estimation methods for limited dependent variable models. In G. S. Maddala, C. R. Rao, & H. D. Vinod (eds.), *Handbook of Statistics* (pp. 519-543). North-Holland, Amsterdam.
- Hakim, C. (1979). *Occupational Segregation*, Research Paper, No. 9, Department of Employment, HMSO (London).
- Hatzinikolaou, D. (2018). *Σημειώσεις για το μάθημα Στατιστική III [Lecture notes for Statistics III]*. University of Ioannina, Department of Economics, Greece.
- Heckman, J.J. (1974). Shadow price, market wages, and labor supply. *Econometrica*, 42 (4) 679–694.
- Heckman, J.J. (1981a). Heterogeneity and State Dependence. In C.F. Manski, D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge.
- Heckman J.J. (1981b). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge MA.
- Heckman, J.J. (2001). Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture. *Journal of Political Economy*, 109(4), 673-748. doi:10.1086/322086
- Heckman, J.J. and Willis, R. W. (1977). A Beta Logistic Model for the Analysis of Sequential Labor Force Participation of Married Women. *Journal of Political Economy*, 85(1), 27-58.

Hernandez-Quevedo, C., Jones, A.M. and Rice, N. (2005). *Reporting bias and heterogeneity in self-assessed health. Evidence from the British Household Panel Survey*. HEDG Working Paper 05/04, University of York, United Kingdom.

Honoré, B.E. and Kyriazidou, E. (2000). Panel Data Discrete Choice Models with Lagged Dependent Variables. *Econometrica*, 68(4), 839–874. doi:10.1111/1468-0262.00139.

Hsiao, C. (2014). *Analysis of Panel Data* (3rd ed.). Econometric Society Monographs. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139839327

Hyslop, D. R. (1999). State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, 67(6), 1255–1294.

Jones, A. M., Rice, N., & Roberts, J. (2010). Sick of work or too sick to work? Evidence on self-reported health shocks and early retirement from the BHPS. *Economic Modelling*, 27(4), 866-880.

Jones, M. K. (2006). Is there employment discrimination against the disabled? *Economics Letters*, 92(1), 32-37.

Jones, A. M., Rice, N., Bago d'Uva, T., and Balia, S. (2013). *Applied health economics* (2nd ed.). Taylor and Francis.

Jones, A. M., Rice, N., & Zantomio, F. (2016). *Acute Health Shocks and Labour Market Outcomes*. University Ca' Foscari of Venice, Dept. of Economics.

Keane, M. (1994). A Computationally Practical Simulation Estimator for Panel Data. *Econometrica*, 62(1), 95-116.

Keane, M.P. and Sauer, R.M. (2006). *Classification Error in Dynamic Discrete Choice Models: Implications for Female Labor Supply Behavior*. IZA Discussion Paper No. 2332.

Keane, M.P. and Sauer, R.M. (2009). Classification Error in Dynamic Discrete Choice Models: Implications for Female Labor Supply Behavior. *Econometrica*, 77(3), pp. 975–991.

Keane M. P., Todd P. E. and Wolpin K. I. (2011). The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications. *Handbook of Labor Economics*, 4A, 331-461.

Keane, M., & Wolpin, K. (2002). Estimating Welfare Effects Consistent with Forward-Looking Behavior. Part II: Empirical Results. *The Journal of Human Resources*, 37(3), 600-622. doi:10.2307/3069683

Keane, M., & Wolpin, K. (2010). THE ROLE OF LABOR AND MARRIAGE MARKETS, PREFERENCE HETEROGENEITY, AND THE WELFARE SYSTEM IN THE LIFE CYCLE DECISIONS OF BLACK, HISPANIC, AND WHITE WOMEN. *International Economic Review*, 51(3), 851-892. Retrieved from <http://www.jstor.org/stable/40784808>

Killingsworth, M. R. (1983). *Labor supply*. New York: Cambridge University Press.

- Killingsworth, M., Heckman, J. (1986). Female labor supply: A survey. In Ashenfelter, O., Layard, R. (Eds.), *Handbook of Labor Economics* (vol. 1, pp. 103–204). North-Holland, Amsterdam.
- Ulrich Kohler, U. and Kreuter, F. 2009. *Data Analysis using Stata* (2nd ed.). Stata Press books, StataCorp LP.
- Laplagne, P., Maurice, G. and Shomos, A. (2007). *Effects of Health and Education on Labour Force Participation*. Productivity Commission Staff Working Paper.
- Lenhart, O. (2019). The effects of health shocks on labor market outcomes: evidence from UK panel data. *European Journal of Health Economics*, 20, 83–98.
- Liljequist, D., Elfving B. and Skavberg Roaldsen, K. (2019). Intraclass correlation – A discussion and demonstration of basic features. PLoS ONE 14(7). Retrieved from <https://doi.org/10.1371/journal.pone.0219854>
- Lindgren, B.W. (1976). *Statistical Theory* (3rd ed.). The Macmillan Co., New York.
- Lindeboom, M. and Kerkhofs, M. (2009). Health and work of the elderly: subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics*, 24 (6), 1024-1046.
- Longhi, S., & Nandi, A. (2015). *A practical guide to using panel data*. SAGE Publications, Inc.
- Lucchetti, R. and Pignini, C. (2017). DPB: Dynamic Panel Binary Data Models in gretl. *Journal of Statistical Software*, 79(8), pp. 1 - 33. doi: <http://10.18637/jss.v079.i08>
- Majeed, T., Forder, P. M., Mishra, G., Kendig, H., & Byles, J. E. (2017). Exploring Workforce Participation Patterns and Chronic Diseases Among Middle-Aged Australian Men and Women Over the Life Course. *Journal of Aging and Health*, 29(2), 343–361.
- Maré, D. (2006). *Constructing Consistent Work-life Histories: A guide for users of the British Household Panel Survey*. ISER Working Paper Series 2006-39.
- McCarthy, H. (2016). Social Science and Married Women's Employment in Post-War Britain. *Past & Present*, 233(1), 269–305.
- McFadden, D. (1977). *Quantitative methods for analysing travel behaviour of individuals: Some recent developments*. Cowles Foundation Discussion Paper No. 474, Cowles Foundation for research in economics at Yale University.
- McFadden, D., Bemmaor, A.C., Caro, F.G., Dominitz, J., Jun, B.H., Lewbel, A., Matzkin, R.L., Molinari, F., Schwarz, N., Willis, R.J., and Winter, J.K. (2005). Statistical analysis of choice experiments and surveys. *Marketing Letters*, 16, 183-196.
- Michaud, P.C. and Tatsiramos, K. (2005). *Employment Dynamics of Married Women in Europe*. IZA Discussion Papers 1706, Institute of Labor Economics.
- Michaud, P.C. and Tatsiramos, K. (2008): *Fertility and female employment dynamics in Europe: the effect of using alternative econometric modeling assumptions*, IZA Discussion

Papers, No. 3853, Institute for the Study of Labor (IZA), Bonn, <http://nbn-resolving.de/urn:nbn:de:101:1-20081202126>

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46(1), 69–85.

Myck, M. and Reed, H. (2006). *A review of static and dynamic models of labour supply and labour market transitions*. IFS Working Papers W06/15, Institute for Fiscal Studies. Retrieved from <https://www.ifs.org.uk/publications/3692>.

Nakamura, A. and Nakamura, M. (1985). Dynamic Models of the Labor Force Behavior of Married Women Which Can Be Estimated Using Limited Amounts of Information. *Journal of Econometrics*, 27(3), 273-98.

Oguzoglu U. (2010). Dynamics of Work Limitation and Work in Australia. *Health Economics*, 19(6), 656-669.

Oguzoglu U. (2016). Disability and Multi-state Labour Force Choices with State Dependence. *The Economic Record*, 92 (296), 28-46.

ONS (2020a). Labour Market Statistics Time Series Dataset, LFS: Female unemployment rate (aged 16 to 64, seasonally adjusted). Retrieved from <https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/timeseries/mgsz/lms>

_ (2020b). Labour Market Statistics Time Series Dataset, Male employment rate (aged 16 to 64, seasonally adjusted). Retrieved from <https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/timeseries/mgsy/lms>

_ (2020c). Labour Market Statistics Time Series Datasets, Estimates of employment, unemployment and economic inactivity for the UK. Retrieved from <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandunemployment/bulletins/employmentintheuk/march2020>

_ (2020d) Retail Price Index Dataset. Retrieved from <https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/chaw/mm23>

Pfarr, C., Schmid, A., & Schneider, U. (2012). Reporting heterogeneity in self-assessed health among elderly Europeans. *Health economics review*, 2(1), 21. <https://doi.org/10.1186/2191-1991-2-21>

Roantree, B. and Vira, K. (2018). *The rise and rise of women's employment in the UK*. IFS Briefing Note BN234. The Institute for Fiscal Studies.

Roos, E., Lahelma, E., Saastamoinen, P., and Elstad, J.I. (2005). The association of employment status and family status with health among women and men in four Nordic countries. *Scandinavian Journal of Public Health*, 33(4), 250–260.

Sayli, M. (2018). *Essays on the added worker effect* (PhD thesis). University of Manchester, Manchester, UK. Retrieved from [https://www.research.manchester.ac.uk/portal/en/theses/essays-on-the-added-worker-effect\(34f8f3ca-9bbe-45d2-a201-f325f5113dd8\).html](https://www.research.manchester.ac.uk/portal/en/theses/essays-on-the-added-worker-effect(34f8f3ca-9bbe-45d2-a201-f325f5113dd8).html)

- Shaw, K. (1994). The Persistence of Female Labor Supply: Empirical Evidence and Implications. *The Journal of Human Resources*, 29(2), 348-378. doi:10.2307/146102
- Social Metrics Commission. (2019). *EQUIVALISATION IN POVERTY MEASURES: CAN WE DO BETTER?* London, UK. Retrieved from <https://socialmetricscommission.org.uk/wp-content/uploads/2019/12/SMC-Equivalisation-Report-2020-01-03-Web.pdf>
- StataCorp. (2007). *Stata 1 Base Reference Manual*. College Station, TX: Stata Press
- StataCorp. (2007). *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP..
- Stern, S. (1989). Measuring the Effect of Disability on Labour Force Participation. *The journal of Human Resources*, 24(3), 361–95. doi:10.2307/145819
- Stewart, M. B. (2006). Maximum simulated likelihood estimation of random effects dynamic probit models with autocorrelated errors. *The Stata Journal*, 6(2), 256–272.
- Stewart, M. B. (2007). The interrelated dynamics of unemployment and low-wage employment. *Journal of Applied Econometrics*, 22(3), 511–531.
- Taylor, Marcia Freed (ed). with John Brice, Nick Buck and Elaine Prentice-Lane (2018) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex
- University of Essex. Institute for Social and Economic Research (2010). British Household Panel Survey: Waves 1-18, 1991-2009. [data collection]. 7th Edition. UK Data Service. SN:5151.
- Van Der Klaauw, W. (1996). Female Labour Supply and Marital Status Decisions: A Life-Cycle Model. *The Review of Economic Studies*, 63(2), 199-235. Retrieved from <http://www.jstor.org/stable/2297850>
- Xiaodong Gong (2010). The added worker effect and the discourage worker effect for married women in Australia. IZA Discussion Paper No. 4816.
- Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20(1), 39–54.
- Wooldridge, J.M. (2000). A framework for estimating dynamic, unobserved effects panel data models with possible feedback to future explanatory variables. *Economics Letters*, 68, 245 – 250.
- Wooldridge, J.M. (2009). *Introductory econometrics: a modern approach* (4th ed.). Mason, Ohio, USA: South-Western Cengage Learning.
- Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge, Massachusetts; London, England: The MIT Press.

Appendices

Appendix A.1

Proofs

Appendix A.1.1

Taylor series expansions of the left and right hand sides around $y_t + w_{0t}^*$ and y_t respectively of

$$u(y_t + w_{1t}^*, 1, \Omega_t) - u(y_t + w_{0t}^* - \gamma_1, 1, \Omega_t) = u(y_t, 0, \Omega_t) - u(y_t - \gamma_1, 0, \Omega_t)$$

give the following connection between the two reservation wages w_{0t}^* and w_{1t}^* :

$$\begin{aligned} & u(y_t + w_{0t}^*, 1, \Omega_t) - u(y_t + w_{0t}^*, 1, \Omega_t) + u_1(y_t + w_{0t}^*, 1, \Omega_t)(w_{1t}^* - w_{0t}^*) - u_1(y_t \\ & \quad + w_{0t}^*, 1, \Omega_t)[y_t + w_{0t}^* - \gamma_1 - (y_t + w_{0t}^*)] \\ & \approx u(y_t, 0, \Omega_t) - u(y_t, 0, \Omega_t) + u_1(y_t, 0, \Omega_t)(y_t - y_t) - u_1(y_t, 0, \Omega_t)(y_t - \gamma_1 - y_t), \end{aligned}$$

where $u_1(\cdot)$ is the partial derivative of the utility function with respect to consumption and the quantity in the brackets $[y_t + w_{0t}^* - \gamma_1 - (y_t + w_{0t}^*)]$ equals $-\gamma_1$.

We divide the above expression by $u_1(y_t + w_{0t}^*, 1, \Omega_t)$ to get

$$w_{1t}^* - w_{0t}^* \approx -\gamma_1 + \gamma_1 \frac{u_1(y_t, 0, \Omega_t)}{u_1(y_t + w_{0t}^*, 1, \Omega_t)}.$$

Therefore, $w_{1t}^* \approx w_{0t}^* - \gamma_1 \left[1 - \frac{u_1(y_t, 0, \Omega_t)}{u_1(y_t + w_{0t}^*, 1, \Omega_t)} \right]$.

From this expression, assuming concave utility (i.e., decreasing marginal utility of consumption), it follows that if $u_1(y_t, 0, \Omega_t) > u_1(y_t + w_{0t}^*, 1, \Omega_t)$. Then the term in the brackets is negative and its product with $-\gamma_1$ is positive.

Therefore,

$$w_{1t}^* \approx w_{0t}^* - \gamma,$$

where

$$\gamma = \gamma_1 \left[1 - \frac{u'(m_t, 0, \Omega_t)}{u'(m_t + w_{0t}^*, 1, \Omega_t)} \right],$$

implies that $w_{1t}^* > w_{0t}^*$.

Appendix A.1.2

Proof of the following equation, taken from Lucchetti and Pignini, p. 5 (see below equation (3.17):

(1)

$$\mathcal{L}_i(\psi) = \int_{-\infty}^{\infty} \Phi[(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i)(2y_{i1} - 1)] \prod_{t=2}^T \Phi[(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)(2y_{it} - 1)] d\Phi\left(\frac{a_i}{\sigma_a}\right),$$

Proof [as in Heckman's (1981) "The Incidental Parameters Problem ...," p. 181, equation (4.2)]:

First, consider $t = 1$. In equation (3.20) let $w = \mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i$ and assume that ε_{i1} is distributed as $N(0, 1)$. Then,

(2)

$$Pr(y_{i1} = 1 | z_{i1}, a_i) = Pr(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i + \varepsilon_{i1} \geq 0) = Pr(\varepsilon_{i1} \geq -w) = Pr(\varepsilon_{i1} < w) = \Phi(w)$$

and

(3)

$$Pr(y_{i1} = 0 | z_{i1}, a_i) = Pr(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i + \varepsilon_{i1} < 0) = Pr(\varepsilon_{i1} < -w) = 1 - \Phi(-w).$$

Combining the last two expressions, we get

$$(4) \quad Pr(y_{i1} / z_{i1}, a_i) = \Phi\{w(2y_{i1} - 1)\} = \Phi\{(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i)(2y_{i1} - 1)\}.$$

Indeed, if $y_{i1} = 1$, (4) reduces to (2); and if $y_{i1} = 0$, (4) reduces to (3).

Next, consider $t \geq 2$. In equation (3.15) let $w = \gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i$ and assume that ε_{it} is $N(0, 1)$. Then, using the same argument as above, we get

$$(5) \quad Pr(y_{it} / y_{i,t-1}, x_{it}, a_i) = \Phi\{(\gamma y_{i,t-1} + x_{it}^T \boldsymbol{\beta} + a_i)(2y_{it} - 1)\}.$$

Assuming independence of $y_{i1}, \dots, y_{iT}$ and letting $\boldsymbol{\psi} = [\boldsymbol{\beta}^T, \gamma, \boldsymbol{\pi}^T, \theta, \sigma_\alpha]$, we see from (4) and (5) that the likelihood function for the i -th individual is

$$(6) \quad \mathcal{J}_i(\boldsymbol{\psi}, a_i) = \Phi\{(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i)(2y_{i1} - 1)\} \prod_{t=2}^T \Phi\{(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)(2y_{it} - 1)\}.$$

Assuming that $a_i | \mathbf{X}_i \sim N(0, \sigma_\alpha^2)$, and hence the ratio a_i/σ_α , conditional on $\mathbf{X}_i$, is distributed as $N(0, 1)$, we can integrate out a_i from (6):

$$\mathcal{J}_i(\boldsymbol{\psi}) = \int_{-\infty}^{\infty} \Phi\{(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i)(2y_{i1} - 1)\} \prod_{t=2}^T \Phi\{(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)(2y_{it} - 1)\} \frac{1}{\sigma_\alpha} \phi\left(\frac{a_i}{\sigma_\alpha}\right) da_i,$$

where ϕ denotes the $N(0, 1)$ density function, or

$$(7) \quad \mathcal{J}_i(\boldsymbol{\psi}) = \int_{-\infty}^{\infty} \Phi\{(\mathbf{z}_{i1}^T \boldsymbol{\pi} + \theta a_i)(2y_{i1} - 1)\} \prod_{t=2}^T \Phi\{(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)(2y_{it} - 1)\} d\Phi\left(\frac{a_i}{\sigma_\alpha}\right),$$

where Φ denotes the cumulative density function (c.d.f.) of the standard normal distribution, $N(0, 1)$. Thus, the proof is complete.

Appendix A.1.3

Proof of equation (3.23), [taken from Bartolucci and Pigini, 2017, equation (1)]:

$$p(y_{it} | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) = \frac{\exp[y_{it}(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}{1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)}.$$

Proof: The logistic distribution is symmetric, bell-shaped (see Greene, 6th Ed., top of p. 774, and Google. The logistic c.d.f. is $\Lambda(z) = e^z/(1 + e^z)$, whereas its probability density function (p.d.f.) is $\lambda(z) = \Lambda'(z) = e^z/(1 + e^z)^2$.

In the static model, we have $y_{it} = \mathbf{1}\{\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i + \varepsilon_{it} \geq 0\}$, $i = 1, \dots, n$, $t = 1, \dots, T$. Assume that ε_{it} has a logistic distribution and let $w = \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i$. Then,

(1)

$$\begin{aligned} Pr(y_{it} = 1 | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) &= Pr(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i + \varepsilon_{it} \geq 0) = Pr(w + \varepsilon_{it} \geq 0) = \\ &= Pr(\varepsilon_{it} \geq -w) = Pr(\varepsilon_{it} \leq w) = \Lambda(w) = ew/(1 + ew), \end{aligned}$$

and

(2)

$$\begin{aligned} Pr(y_{it} = 0 | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) &= Pr(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i + \varepsilon_{it} \leq 0) = Pr(w + \varepsilon_{it} \leq 0) = Pr(\varepsilon_{it} \leq -w) \\ &= 1 - Pr(\varepsilon_{it} \geq w) = 1 - \Lambda(w) = 1 - ew/(1 + ew) = 1/(1 + ew). \end{aligned}$$

Combining (1) and (2), we get

$$(3) \quad Pr(y_{it} | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) = \exp(y_{it}w)/(1 + ew).$$

Indeed, if $y_{it} = 1$, (3) reduces to (1); and if $y_{it} = 0$, (3) reduces to (2). Putting back $w = \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i$ gives equation (3.23)

$$(4) \quad Pr(y_{it} | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) = \exp[y_{it}(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]/[1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)].$$

Appendix A.1.4

Proof of equation (3.24), [taken from Bartolucci and Pigini, 2017, p. 3]:

$$p(\mathbf{y}_i | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) = \frac{\exp(\sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta}) \exp(a_i y_{i+})}{\Pi_t [1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}$$

If $y_{i1}, \dots, y_{iT}$ is indeed a random sample, then $y_{i1}, \dots, y_{iT}$ are independent, so their joint probability [using equation (4) from the previous proof] is

$$\begin{aligned} p(\mathbf{y}_i | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) &= p(y_{i1} | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) \dots p(y_{iT} | \mathbf{X}_i, \alpha_i, \boldsymbol{\beta}) \\ &= \frac{\exp[y_{i1}(\mathbf{x}_{i1}^T \boldsymbol{\beta} + a_i)]}{1 + \exp(\mathbf{x}_{i1}^T \boldsymbol{\beta} + a_i)} \dots \frac{\exp[y_{iT}(\mathbf{x}_{iT}^T \boldsymbol{\beta} + a_i)]}{1 + \exp(\mathbf{x}_{iT}^T \boldsymbol{\beta} + a_i)} = \frac{\exp[\sum_t y_{it}(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}{\Pi_t [1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]} \\ &= \frac{\exp(\sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta}) \exp(a_i y_{i+})}{\Pi_t [1 + \exp(\mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}, \end{aligned}$$

where $y_{i+} = \sum_t y_{it}$. Thus, the proof is complete.

Appendix A.1.5

Proof of equation (3.25), [taken from Bartolucci and Pigini, 2017, p. 3]:

$$(1) \quad p(\mathbf{y}_i | \mathbf{X}_i, y_{i+}; \boldsymbol{\psi}) = \frac{p(\mathbf{y}_i | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})}{p(y_{i+} | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})},$$

where the statistic $y_{i+} = \sum_t y_{it}$ (the total score) is sufficient for the parameter α_i . Here, we assume that $\mathbf{y}_i = (y_1, \dots, y_T)^T$ is a $T \times 1$ vector of discrete random variables, so the function p is actually a probability (denoted as Pr), and the symbols p and Pr can be used interchangeably here.

Proof: Let the sufficient statistic be denoted as $T = \sum_t y_{it} = y_{i+}$; let k be a possible value of T ; and let $A(k)$ be the set of all points $y_1, \dots, y_T$, such that $T = k$. By the definition of a sufficient statistic, for any given value of α_i in the parameter space and a matrix of explanatory variables $\mathbf{X}_i$, the conditional distribution of $\mathbf{y}_i = (y_1, \dots, y_T)^T$ given $T = k$ does not depend on α_i . That is,

$$(2) \quad Pr(\mathbf{y}_i | \mathbf{X}_i, \alpha_i, T = k; \boldsymbol{\psi}) = \frac{Pr(\mathbf{y}_1, \dots, \mathbf{y}_T \text{ and } T=k | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})}{Pr(T=k | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})} = Pr(\mathbf{y}_i | \mathbf{X}_i, T; \boldsymbol{\psi}).$$

The numerator and the denominator of (2) are nonzero only when $T = k$. In that case, however, the condition $T = k$ is redundant in the numerator of (2), i.e., it adds no information, since T is a function of $y_1, \dots, y_T$, which are already present, so it can be dropped. Thus, (2) can be written as (see Degroot, 1986, p. 359; Lindgren, 1976, p. 227; and Μπερτσεκός and Τσιτσικλής, 2016, p. 591)

$$(3) \quad \frac{Pr(\mathbf{y}_i | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})}{Pr(\mathbf{y}_{i+} | \mathbf{X}_i, \alpha_i; \boldsymbol{\psi})} = Pr(\mathbf{y}_i | \mathbf{X}_i, \mathbf{y}_{i+}; \boldsymbol{\psi}).$$

Since p and Pr are used interchangeably here, equation (3) is just another way of writing (1).

Appendix A.1.6

Proof of equation (3.30), [taken from Lucchetti and Pigni, 2017, p. 6]:

For the dynamic logit model, we have that

$$p(y_{it} | y_{i1}, \dots, y_{i,t-1}, \mathbf{X}_i, \alpha_i; \boldsymbol{\psi}) = \frac{\exp[y_{it}(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}{1 + \exp(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)}$$

As is noted in the thesis, equation (3.30) is an extension of (3.23) for the static logit model, which was proved above. In particular, it augments (3.23) to include $y_{i,t-1}$ among the explanatory variables.

Appendix A.1.7

Proof of equation (3.31), [this is equation (3) of Bartolucci and Nigro, 2010, p. 721, or equation (6) of Luccetti and Pignini, 2017, p. 7]:

The joint probability of the overall vector of the response variables, $\mathbf{y}_i$, conditional on $\mathbf{X}_i$, α_i , and y_{i1} , is

$$p(\mathbf{y}_i | \mathbf{X}_i, \alpha_i, y_{i1}; \boldsymbol{\psi}) = \frac{\exp(y_{i*}\gamma + \sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta} + y_{i+} a_i)}{\prod_t [1 + \exp(y_{i,t-1}\gamma + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]},$$

where $y_* = \sum_t y_{i,t-1} y_{it}$, $y_+ = \sum_t y_{it}$, $\boldsymbol{\psi} = (\boldsymbol{\beta}, \gamma)$, and sums and products go from $t = 2$ to T .

Proof: To economize on space, let $W = (\mathbf{X}_i, \alpha_i, \boldsymbol{\psi})$. Evidently, since in this case y_{it} depends on $y_{i,t-1}$, the observations $y_{i1}, \dots, y_{iT}$ are not independent, so the joint probability $Pr(y_{i2}, \dots, y_{iT} | W, y_{i1})$ is *not* equal to the product of the marginal probabilities, $Pr(y_{i2} | W, y_{i1}) \dots Pr(y_{iT} | W, y_{i1})$. Using the product law for conditional densities (see Wooldridge, 2010, p. 524), however, we have that

$$\begin{aligned} & Pr(y_{i2}, \dots, y_{iT} | W, y_{i1}) \\ &= Pr(y_{i3}, \dots, y_{iT} | W, y_{i1}, y_{i2}) Pr(y_{i2} | W, y_{i1}) \\ &= Pr(y_{i4}, \dots, y_{iT} | W, y_{i1}, y_{i2}, y_{i3}) Pr(y_{i3} | W, y_{i1}, y_{i2}) Pr(y_{i2} | W, y_{i1}) \\ &= \dots \\ &= \\ & Pr(y_{i2} | W, y_{i1}) Pr(y_{i3} | W, y_{i1}, y_{i2}) \dots Pr(y_{iT-1} | W, y_{i1}, \dots, y_{iT-2}) Pr(y_{iT} | W, y_{i1}, \dots, y_{iT-1}). \end{aligned}$$

Substituting from equation (3.30), which has just been proved, this probability becomes

$$\begin{aligned}
p(y_{i2}, \dots, y_{iT} | W, y_{i1}) &= \frac{\exp[y_{i2}(\gamma y_{i1} + \mathbf{x}_{i2}^T \boldsymbol{\beta} + a_i)]}{1 + \exp(\gamma y_{i1} + \mathbf{x}_{i2}^T \boldsymbol{\beta} + a_i)} = \dots \\
&= \frac{\exp[y_{iT}(\gamma y_{i,T-1} + \mathbf{x}_{iT}^T \boldsymbol{\beta} + a_i)]}{1 + \exp(\gamma y_{i,T-1} + \mathbf{x}_{iT}^T \boldsymbol{\beta} + a_i)} \\
&= \frac{\exp(\gamma \sum_t y_{i,t-1} y_{it} + \sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i \sum_t y_{it})}{\prod_t [1 + \exp(y_{i,t-1} \gamma + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]} \\
&= \frac{\exp(\gamma y_{i*} + \sum_t y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i y_{i+})}{\prod_t [1 + \exp(y_{i,t-1} \gamma + \mathbf{x}_{it}^T \boldsymbol{\beta} + a_i)]}.
\end{aligned}$$

Thus, the proof is complete.

Appendix A.1.8

Proof of equation (3.33), [this is equation (8) of Lucchetti and Pignini, 2017, p. 7]:

(1)

$$\begin{aligned}
p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, y_{i+}; \boldsymbol{\psi}) &= \frac{p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi})}{p(y_{i+} | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi})} \\
&= \frac{\exp[\sum_{t=2} y_{it} y_{i,t-1} \gamma + \sum_{t=2} y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta}_1 + y_{iT}(\mu + \mathbf{x}_{iT}^T \boldsymbol{\beta}_2)]}{\sum_{\mathbf{b}: \mathbf{b}_+ = y_{i+}} \exp[\sum_{t=2} b_t b_{t-1} \gamma + \sum_{t=2} b_t \mathbf{x}_{it}^T \boldsymbol{\beta}_1 + b_T(\mu + \mathbf{x}_{iT}^T \boldsymbol{\beta}_2)]},
\end{aligned}$$

where the sum $\mathbf{b}_+ \equiv \sum_t b_t$ ranges over all possible binary response T -vectors containing zeros and ones $\mathbf{b} = (b_1, \dots, b_T)^T$ in a set $\mathbb{B} \equiv \{\mathbf{b}: \mathbf{b} \in \{0,1\}^T\}$.

Proof: The first equal sign in (1) is proved as in the case of equation (3.25), so let's prove the second equal sign in (1). Lucchetti and Pignini's (2017, p. 7) equation (7) is the *assumption* or *definition* of Bartolucci and Nigro (2010, p. 723, equation (4)), namely,

(2)

$$\begin{aligned}
&p(\mathbf{y}_i | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi}) \\
&= \frac{\exp[\sum_{t=2} y_{it} y_{i,t-1} \gamma + \sum_{t=2} y_{it} \mathbf{x}_{it}^T \boldsymbol{\beta}_1 + y_{iT}(\mu + \mathbf{x}_{iT}^T \boldsymbol{\beta}_2) + y_{i+} \alpha_i]}{\omega(\mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi})},
\end{aligned}$$

where

(3)

$$\omega(\mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi}) = \sum_{\mathbf{b} \in \mathbb{B}} \exp\left[\sum_{t=2} b_t b_{t-1} \gamma + \sum_{t=2} b_t \mathbf{x}_{it}^T \boldsymbol{\beta}_1 + b_T (\mu + \mathbf{x}_{iT}^T \boldsymbol{\beta}_2) + b_+ \alpha_i\right]$$

is a normalizing constant that does not depend on $\mathbf{y}_i = (y_{i2}, \dots, y_{iT})$; see Bartolucci and Nigro (2010, p. 723). Bartolucci and Nigro (2010, p. 727) show that, under definition (2), we have that

(4)

$$\begin{aligned} p(y_{i+} | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi}) &= \sum_{\mathbf{b}(y_{i+})} p(\mathbf{y}_i = \mathbf{b} | \mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi}) \\ &= \frac{\exp(y_{i+} \alpha_i)}{\omega(\mathbf{X}_i, y_{i1}, \alpha_i; \boldsymbol{\psi})} \sum_{\mathbf{b}(y_{i+})} \exp\left[\sum_{t=2} b_t \mathbf{x}_{it}^T \boldsymbol{\beta}_1 + b_T (\mu + \mathbf{x}_{iT}^T \boldsymbol{\beta}_2) + b_{i*} \gamma\right], \end{aligned}$$

where the sum $\sum_{\mathbf{b}(y_{i+})}$ is restricted to all response configurations $\mathbf{b}$ such that $b_+ = y_{i+}$ and $b_{i*} = y_{i1} b_1 + \sum_{t>2} b_t b_{t-1}$. Now put (2) in the numerator of the first term on the right-hand side of (1) and (4) in the denominator. Recognizing that $\sum_{t=2} b_t b_{t-1} = b_{i*}$, the result is equation (1), and the proof is complete.

Appendix A.1.9

Derivation of the likelihood function in the case of Mundlak's (1978) correlated random effects model; see section 3.3.3.

Consider the following dynamic binary response model:

$$y_{it} = 1 \{y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i^T \mathbf{c} + \omega_1 y_{i1} + \xi_i + \varepsilon_{it} \geq 0\}, \quad t = 2, \dots, T,$$

where ξ_i is i.i.d. $\sim N(0, \sigma_\xi^2)$ and independent of $\mathbf{x}_{it}$ and ε_{it} for all i, t , and where $\varepsilon_{it} \sim \text{i.i.d.} N(0, 1)$. Thus, $y_{i,t}$ given $(y_{i,t-1}, \dots, y_{i1}, \mathbf{X}_i, \omega_1, \gamma, \boldsymbol{\beta}, \mathbf{c}, \xi_i)$ follows a probit model. Let $W = (\omega_1, \mathbf{x}_{it}, \bar{\mathbf{x}}_i, \gamma, \boldsymbol{\beta}, \mathbf{c}, \xi_i)$.

Evidently, since in this case y_{it} depends on $y_{i,t-1}$, the observations $y_{i1}, \dots, y_{iT}$ are not independent, so the joint probability $Pr(y_{i2}, \dots, y_{iT} | W, y_{i1})$ is *not* equal to the product of the marginal probabilities, $Pr(y_{i2} | W, y_{i1}) \dots Pr(y_{iT} | W, y_{i1})$. Using the product law for conditional densities (see Wooldridge, 2010, p. 524), however, we have that

$$\begin{aligned}
& Pr(y_{i2}, \dots, y_{iT} | W, y_{i1}) \\
&= Pr(y_{i3}, \dots, y_{iT} | W, y_{i1}, y_{i2}) Pr(y_{i2} | W, y_{i1}) \\
&= Pr(y_{i4}, \dots, y_{iT} | W, y_{i1}, y_{i2}, y_{i3}) Pr(y_{i3} | W, y_{i1}, y_{i2}) Pr(y_{i2} | W, y_{i1}) \\
&= \dots \\
&= \\
& Pr(y_{i2} | W, y_{i1}) Pr(y_{i3} | W, y_{i1}, y_{i2}) \dots Pr(y_{iT-1} | W, y_{i1}, \dots, y_{iT-2}) Pr(y_{iT} | W, y_{i1}, \dots, y_{iT-1}).
\end{aligned}$$

Since the random variables $y_{i2}, \dots, y_{iT}$ are discrete, this equation can also be written as a product of conditional probability functions, namely,

$$\begin{aligned}
& f(y_{i2}, \dots, y_{iT} | W, y_{i1}) \\
&= \\
& f(y_{i2} | W, y_{i1}) f(y_{i3} | W, y_{i1}, y_{i2}) \dots f(y_{iT-1} | W, y_{i1}, \dots, y_{iT-2}) f(y_{iT} | W, y_{i1}, \dots, y_{iT-1}).
\end{aligned}$$

But

$$\begin{aligned}
Pr(y_{it} = 1 | W, y_{i1}) &= Pr(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i + \varepsilon_{it} \geq 0) \\
&= Pr(\varepsilon_{it} \geq -\gamma y_{i,t-1} - \mathbf{x}_{it}^T \boldsymbol{\beta} - \bar{\mathbf{x}}_i \mathbf{c} - \omega_1 y_{i1} - \xi_i) \\
&= Pr(\varepsilon_{it} \leq \gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i) \\
&= \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i),
\end{aligned}$$

because ε_{it} is i.i.d. $N(0,1)$, where Φ denotes the cumulative density function of the $N(0, 1)$ distribution. Similarly,

$$\begin{aligned} Pr(y_{it} = 0 \mid W, y_{i1}) &= Pr(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i + \varepsilon_{it} < 0) \\ &= Pr(\varepsilon_{it} < -\gamma y_{i,t-1} - \mathbf{x}_{it}^T \boldsymbol{\beta} - \bar{\mathbf{x}}_i \mathbf{c} - \omega_1 y_{i1} - \xi_i) \\ &= 1 - \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i). \end{aligned}$$

Thus,

$$\begin{aligned} f(y_{it} \mid W, y_{i1}) \\ = [\Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i)]^{y_{it}} [1 - \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i)]^{1-y_{it}}. \end{aligned}$$

Hence, the likelihood function is

$$\begin{aligned} f(y_{i2}, \dots, y_{iT}) \mid W, y_{i1} \\ = \prod_{t=2}^T \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i)]^{y_{it}} [1 \\ - \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i)]^{1-y_{it}} \end{aligned}$$

Following Wooldridge (2010, pp.613 and pp.627-628), we integrate out ξ_i , to obtain

$$\begin{aligned} f(y_{i2}, \dots, y_{iT}) \mid \omega_1, \mathbf{x}_{it}, \bar{\mathbf{x}}_i, \gamma, \boldsymbol{\beta}, \mathbf{c}, \sigma_\xi^2, y_{i1} \\ = \int_{-\infty}^{\infty} \left[\prod_{t=2}^T \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i)]^{y_{it}} [1 \right. \\ \left. - \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{c} + \omega_1 y_{i1} + \xi_i)]^{1-y_{it}} \right] \frac{1}{\sigma_\xi} \Phi\left(\frac{\xi_i}{\sigma_\xi}\right) d\xi_i \end{aligned}$$

or

$$f(y_{i2}, \dots, y_{iT}) \mid \omega_1, \mathbf{x}_{it}, \bar{\mathbf{x}}_i, \gamma, \boldsymbol{\beta}, \mathbf{a}, \sigma_\xi^2, y_{i1}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left[\prod_{t=2}^T \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{a} + \omega_1 y_{i1} + \xi_i) \right]^{y_{it}} [1 \\
&\quad - \Phi(\gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + \bar{\mathbf{x}}_i \mathbf{a} + \omega_1 y_{i1} + \xi_i)]^{1-y_{it}} \Big] d\Phi\left(\frac{\xi_i}{\sigma_\xi}\right).
\end{aligned}$$

Appendix A.2

Correlation Matrix

	Y	HLPRBA	HLPRBB	HLPRBC	HLPRBD	HLPRBE	HLPRBF	HLPRBG	HLPRBH	HLPRBI
Y	1.0000									
HLPRBA	-0.1590	1.0000								
HLPRBB	-0.0770	0.1253	1.0000							
HLPRBC	0.0450	0.0361	0.0414	1.0000						
HLPRBD	-0.0096	0.0453	0.0157	0.0200	1.0000					
HLPRBE	-0.1042	0.1080	0.0831	0.0250	0.1293	1.0000				
HLPRBF	-0.0775	0.1544	0.0897	0.0289	0.0430	0.1119	1.0000			
HLPRBG	-0.0371	0.0946	0.0214	0.0462	-0.0088	0.0208	0.0613	1.0000		
HLPRBH	-0.0642	0.0256	-0.0060	0.0322	-0.0126	-0.0103	0.1417	0.0769	1.0000	
HLPRBI	-0.1906	0.1273	0.0710	0.0195	-0.0044	0.0711	0.0735	0.0778	0.0308	1.0000

Appendix A.3

Consumer Price Index

Year	Base year: January 1987=100
1991	133.5
1992	138.5
1993	140.7
1994	144.1
1995	149.1
1996	152.7
1997	157.5
1998	162.9
1999	165.4
2000	170.3
2001	173.3
2002	176.2

Source: Office for National Statistics – Consumer price inflation

Appendix A.4

McClements Equivalence Scale

Household Size and Composition	After Housing Costs
Head	0.55
Spouse	0.45
Other second adult	0.45
Third adult	0.45
Further adult	0.40
Dependent child aged:	
0-1	0.07
2-4	0.18
5-7	0.21
8-10	0.23
11-12	0.26
13-15	0.28
16+	0.38

Source: British Household Panel Survey – User Manual, Volume A

Appendix A.5

Variable name in BHPS	Definition	Construction	How it appears in the analysis
wJBSTAT	Current labour force status: “Please look at this card and tell me which best describes your current situation?” The interviewer shows a card that includes the following possible options: self-employed, in paid-employment (full-time or part-time), unemployed, retired, looking after a family, full-time student/at school, long-term sick or disabled, on maternity leave, on government training scheme, something else	We created two categories of employment state for the respondents. One category, named “employed,” contains respondents who are self-employed and in paid employment, either full-time or part-time. The second category, named “non-employed,” contains respondents who are unemployed and, additionally, the group of respondents who are family carers, full-time students/at school and on maternity leave. We constructed a binary variable y that equals 1 if the woman works and 0 if she does not work. The reference category is $y = 0$.	In chapter 4, the variable appears as “Employment status” and in chapter 5, the variable appears as “participation”.
wHLLT	“Does your health in any way limit your daily activities compared to most people of your age?”	We constructed a binary indicator whether the woman reports having limiting health problems. The reference category is “no limiting health problems.”	Limiting health problems
wHLSTAT	“Please think back over the last 12 months about how your health has been. Compared to people of your own age,	We recoded self-assessed health as a four-category scale; excellent, good or very good, fair and poor or very poor.	Health status

	would you say that your health has on the whole been excellent/good/fair/poor/very poor?"	The reference category is "poor or very poor." We also constructed binary indicators for each health state.	Excellent health Very good or good health Fair health Poor or very poor health HLSTAT01
wHLPRBA	The respondent is asked whether she has any of a list specific health problems related to arms, legs, hands, etc.	We constructed a binary dummy variable for the presence or not of the specific health problem.	Arms, legs, hands
wHLPRBI	The respondent is asked whether she has any of a list specific health problems related to anxiety/depression.	We constructed a binary dummy variable for the presence or not of the specific health problem.	Anxiety/Depression
wBLESLEN	In BHPS, the employment status history is present at wave 2 only in months.	We converted the months to years and for each of the subsequent waves, we added an additional year of work experience if the respondent stated "employed", that is $y = 1$. We did not add an additional year if she stated "non-employed", $y = 0$.	Years of experience
wQFACHI	Highest academic qualification	We converted the educational categories to years of education. Years of education are imputed from the categorical scheme: "HIGHEST DEGREE" = 18 years, "1 st DEGREE" = 16 years, "HND/Teaching O/CSE" = 15 years, "A LEVEL" = 13 years, "O LEVEL/CSE" = 11 years.	Years of education

		For the last category “OTHER QUALIFICATIONS” we allowed for another variable that indicates schooling leaving age (wSCEND). If the respondent claims other qualifications and her schooling leaving age is greater than 16 years, we consider that she has completed 11 years of education. If the respondent claims other qualifications and her schooling leaving age is less than 16 years, we consider that she has completed the ‘schooling leaving age’ years minus the lowest compulsory five years of education.	
1) wNCH02 2) wNCH34 3) wNCH511	1) Number children in household aged 0-2 2) Number children in household aged 3-4 3) Number children in household aged 5-11	The presence of children living in the household at different ages is included in the analysis as a binary indicator of whether the woman has any children in a specific age group of children or not.	1) Presence of children aged 0-2 years, 2) Presence of children aged 3-4 years, 3) Presence of children aged 5-11 years
wFIHHMN on record wHHRESP and wFIMNL on record wINDRESP	1) Household income: month before interview	We created a transitory non-labour income variable as the deviation from the mean non-labour income (over the twelve years). In BHPS, monthly permanent non-labour income is derived from monthly household income. Then we subtracted the	Transitory Income

	2) Labour income: last month	monthly woman's labour income from monthly household income and we got the non-labour income, equivalised and deflated to 1991 sterling pounds values. We used the natural logarithm to measure changes in transitory non-labour income.	
wFIHHMN on record wHHRESP and wFIMNL on record wINDRESP	1) Household income: month before interview 2) Labour income: last month	We created the mean (over the twelve years) of non-labour income. In BHPS, monthly permanent non-labour income is derived from monthly household income. Then we subtracted the monthly woman's labour income from monthly household income and we got the non-labour income, equivalised and deflated to 1991 sterling pounds values. Finally, we used the natural logarithm to measure changes in permanent non-labour income.	Permanent Income
wSPJOB	Whether spouse/partner employed now	We constructed a dummy variable that identifies whether the spouse or partner is unemployed. The reference category is "partner is employed."	Spouse unemployed
wAGE	Age at Date of Interview	unchanged	age

Appendix A.6

Record Type	Description	Purpose
HHRESP	Contains data from the Household Questionnaire for sample selected <u>households</u>	Gives information on household income
INDRESP	Contains individual data from main individual questionnaire	Gives information on personal characteristics
XWAVEID	Contains information for matching individuals between waves	To merge information for the same respondent in successive waves

Appendix A.7

The following table contains the information on the pattern of missing values for 8 variables in our final dataset. The pattern of missing values is described by using a binary indicator which takes on the value of 0 if there are missing values and 1 otherwise. From the first line we see that for 3,865 of women, i.e., 97 percent of the total number, we have data on all variables. From the second line onward, the first column lists the number of missing values and the second lists the percent missing for all 8 variables together. For example, from the second line we see that there are 63 women with missing values, the second column indicates that up to 2 percent of the observations are missing, and the distribution shows that the variable “employment state” has missing values. The remaining lines indicate that only a very low percentage of values are missing (less than 1 percent).

Missing-value patterns

Frequency	Percent	Missing-value pattern							
		Spouse unemployed	Limiting Health Problems	Health Status	Arms, legs, hands, etc.	Anxiety/Depression	Years of Education	Transitory income	Employment State
3865	97%	1	1	1	1	1	1	1	1
63	2	1	1	1	1	1	1	1	0
33	<1	1	1	1	1	1	1	0	1
10	<1	1	1	1	1	1	0	1	1

4	<1	1	1	1	0	0	1	1	1
2	<1	1	1	0	1	1	1	0	1
2	<1	1	1	0	1	1	1	1	1
2	<1	1	1	1	1	1	0	0	1
1	<1	0	1	1	1	1	1	1	1
1	<1	1	0	0	0	0	1	1	1
1	<1	1	0	1	0	0	1	1	1
3.984									

