Author/s:
Mai, Duy Tung

Title:
An Econometric Analysis of Multivariate Fractional Time Shares Using Australian Longitudinal Household Data

Date:
2020

An Econometric Analysis of Multivariate Fractional Time Shares Using Australian Longitudinal Household Data

Tung Duy Mai

Submitted in partial fulfilment of the requirements of the degree
Doctor of Philosophy (with coursework component).

January, 2021

Department of Economics
The University of Melbourne

Abstract

Household time use is an important element of labour supply, gender division of labour and the household production function. My dissertation highlights the importance of individual heterogeneity and the natural boundedness of time allocations in time-use research. First, it takes advantage of the rich longitudinal data from the Household, Income and Labour Dynamics in Australia survey,such data are rare in time-use literature. It develops a novel class of estimators that are capable of accommodating more econometric features of the study than other existing methods (including individual fixed effects, non-linearity, and the multivariate nature of such joint decisions) and applies it to the empirical study of impact of children onto couples' time uses. Second, it extends the econometric framework to adopt random coefficients feature. Finally, it studies the factors driving the joint allocation of time use within coupled families: individual preferences, observable changes, and unknown shocks. Overall, I find that (i) the negative impacts of parenthood on women's work hours have been narrowed over time; (ii) the extension to random coefficients is non-trivial; and (iii) individual preferences play an important role in determining the cooperation of household time allocation.

Declaration

This is to certify that

(i) the thesis comprises only my original work towards the degree Doctor of Philosophy (with coursework component),

(ii) due acknowledgement has been made in the text to all other material used,

(iii) the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Signed

Dated: January 31, 2021.

Acknowledgments

I am truly grateful to my supervisors Kevin Staub and Christopher Skeels for their invaluable guidance and unconditional support.

I am also grateful to the department's Ph.D. Director Kalvinder Shields for her constant care and encouragement for me.

I also thank Rainer Winkelmann, Frank Windmeijer and Daniel Hamermesh for their many useful comments.

All errors are my own.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Time as a resource has seemingly become increasingly scarce and one of the most important constraints for a person. Individuals now have relatively less time compared to other types of capital. From a macro perspective, while real gross domestic product per capita has tripled in rich countries in the past 60 years, the number of minutes in a day has not changed and we do not live a multi-fold longer (Hamermesh (2016)). The longevity has risen merely by a few years over the same period of time. The way each individual uses their time can be vastly different. Depending on own preference, personal characteristics, and the environment one lives in, he or she might feel either time-poor, or time-rich.

A major riddle is how variant people's value of time affects their consumption of time and other resources. In particular, they involve how people choose to allocate their time between work and home. More importantly, in the context of families, how market opportunity costs of its members' time may affect the joint decision of task allocation. There is a cooperative structure of household time allocation that is left in a black box. As time is getting more valuable, household time allocation is increasingly significant at explaining household production function and intra-household bargaining patterns. Not only does it reveal how couples handle impacts from exogenous life events but also explains individual preferences over household time allocations.

Prior to 1960s, household production function was a "black box" to mainstream economics (Heckman (2015)). Mitchell (1912) wrote about the "backward art of spending money" and compared the efficiency of firms in producing goods for the market to the inefficiency of households in producing domestic services. Kuznets (1934) stated that Gross National Products (GNP) accounts omitted important components of household production. Reid (1934) wrote a textbook on home production aimed at students of home economics. It offered practical advice, sketched some analytical principles and interesting speculation about the future of households and the role of women. However, it was not until the work of Mincer (1962) and Becker (1965) that economists began to model households formally as an entity engaging in taking inputs and producing outputs. Becker (1965) laid the analytical foundations for the framework of household production and time allocation. Since these seminal works, the interest for studying household time allocation has officially been on the rise.

Household labour participation is driven by partners' value of time. In a two-person household, the prices of each spouses time affect their choices between market and non-market activities, a topic that has been studied in the labour economics literature (Ashenfelter and Heckman (1974)). This broadly differentiates between paid work activities and unpaid home activities. While this line of research has been used to examine how labour supply responds to prices, a broader question is how various components of the non-market time of each spouse respond to their values of time.

In the last decade the important literature on bargaining within household has been expanded beyond the market and non-market distinction. It now also considers how choices are made outside the market (Cherchye, De Rock, and Vermeulen (2012)). Complementaries in family labour supply and leisure within or beyond the household are a key policy issue, as they represent a channel through which reforms targeted at specific segments of the population can ultimately affect a wider set of individuals. When the value of leisure time for an individual depends on the amount of leisure enjoyed by her spouse, reforms of the welfare state, or tax reforms, or changes in work week regulations aimed at some segments of the workforce may impact individual behaviour well beyond the targeted population (Alesina, Glaeser, and Sacerdote (2005)).

The first disaggregation of non-market time is between leisure and unpaid work (Robinson and Godbey (1997), Gershuny (2003)). The usual practice is to ask whether there are market substitutes for an activity. For example, time spent cooking can be seen as unpaid work because restaurant meals are a market substitute. Similarly, a housekeeper can clean the house and a babysitter can provide childcare. In all of these cases, the implicit assumption is the absence of process preferences where some do enjoy cooking or caring for their children. In contrast, I can not hire someone to exercise or to listen to music on my behalf, or to take a rest and sleep for me.

As labour supply models often treat active time that is not allocated to market work as a single category of non-market activities, time use literature recognised that individuals and households allocate time among different types of unpaid work subject to budget, time and a household technology constraints. As childcare time allocation provide insights for the division of labour with regards to who would spend more time with the child and who would take more work hours to meet the increased financial needs, it helps measure the extent to which parents spending time with their children versus any other commitments. The joint structure of time shares between parents is hence useful for the overall understanding of labour distribution during parenthood.

It is important that time-use studies take into account the presence of two aspects: individual heterogeneity and the natural boundedness of time allocations.

First, in terms of individual heterogeneity, the literature has been behind due to a lack of longitudinal data. Hamermesh (2016) pointed out that most of the literature was conducted in a cross-sectional setting (Argyrous and Rahman (2017), Bauer and Sonchak (2017)) which is compromised by potential unobservable individual effects. This limitation mainly results from the unavailability of time-use data in a panel setting, as most collections of time-use diaries have been only in a cross-sectional setting. This lack of longitudinal information causes validation issues for these studies and incentivised researchers to look for alternative data: retrospectively self-reported panel surveys of time use (Craig and Siminski (2011), Foster and Stratton (2018)). Although there are drawbacks due to larger measurement error by its retrospective nature, the advantages of panel time use data outweighs these, especially in this context where correlated individual effects are of high concern.

Second, time allocation is bounded. In a given period such as a day or a week, each person only has a fixed amount of time to allocate to different activities of their choice. For example, the hours one can spend on paid work in a week could range from 0 to a hypothetical $168 = (24 \times 7)$. These outcomes can be normalised to the unit interval of $[0; 1]$ and, importantly, could take the boundary values of 0 and 1. Due to the boundedness of outcome variables, a non-linear model is preferred to a standard linear form. While non-linearity brings the benefits of preserving the bounded nature of the dependent variable, it also increases the complexity of the model. Each type of non-linear function often requires its own transformation technique.

To address these aspects, I take advantage of the rich longitudinal data from the Household, Income and Labour Dynamics in Australia (HILDA) survey (Department of Social Services and Melbourne Institute of Applied Economic and Social Research (2017)) and apply a non-linear model for my econometric framework. While there are a number of different functional forms to fit a fraction, I choose the logistic function

$$\mathbb{E}[y|x] = \Lambda(x) = \frac{e^x}{1 + e^x}$$

as it provides a tractable form of transformation. As my approach concerns the importance of the ability to manipulate the link function, the choice of logistic directly relates to how I transform the conditional expectation, producing the corresponding moment conditions. Beside the choice of logit, probit is another popular link function. A different approach could be applied for the case, however, I observe that the inverse of probit link function will embed more limitations on my ability to flexibly work with the variables.

To the best of my knowledge, there is no current tool that could handle a panel setting data with non-linearity for multinomial fractional outcomes. It thus motivates the development of my own estimator. I build a class of estimators that are capable of accommodating more required features than other existing methods.

1.1 Longitudinal data

The HILDA Survey is a nationally representative longitudinal study of Australian households which commenced in 2001. It is managed by the Melbourne Institute of Applied Economic and Social Research at the University of Melbourne and funded by the Australian Government Department of Social Services. It provides longitudinal data on the lives of Australian residents. Its primary objective is to support research questions falling within three interconnected areas of income, labour and family dynamics. As it is a household-based panel study, it interviews all household members (15 years and over) of the selected households and then re-interviews the same people in subsequent years. This dataset provides a unique insight into the dynamics of Australian labour force. Beside its focus on the labour market outcomes, it covers a wide range of topics such as education, health, fertility, retirement, family relationship, etc. The dataset used in this dissertation is the 16th release of the HILDA survey, incorporating data collected from 2002 through to 2016. It does not include the first wave as the time use module only started from the second wave.

The outcome variable is originated from the answers to the question *"How much time would you spend on each of the following activities in a typical* *week?"*. They include usual hours one spends on nine different activities. They are mutually exclusive, however, not collectively exhaustive. These categories only cover the active hours of doing paid and unpaid works. Leisure, recreational activities, resting, and sleeping time were not in there. Therefore, I selected paid employment, child care, housework out of this list of nine and let the remaining time as the last group. The resulting four categories are paid employment, child care, housework and the rest. They sum up to a maximum total of 168 hours per week. The last category is loosely named leisure which also includes rest, sleeping, voluntary and other recreational personal activities. The four groups are now mutually exclusive and collectively exhaustive of total available hours. I later define my time shares as a proportion of time spent out of this total.

I define $w_{it}^j, ch_{it}^j, h_{it}^j$ and l_{it}^j, where $\sum_{j\in\{m,f\}} w_{it}^j + ch_{it}^j + h_{it}^j + l_{it}^j = 1$, to be the time shares of work, child care, home production, and leisure hours for j being either husband or wife of household i at time t, respectively. For the identification purpose of the multinomial logit setting, it requires one category to be omitted. I opted for the female leisure share l_{it}^f though choices of omitted group would not change the nature of the results. Systematically, the reference group will act as the baseline where all the interpretations of an impact from explanatory variables will be in a relative form. A change in reference group will alter the estimates and their interpretations, however, it would not change the nature of the parameters for average partial effects (APEs).

The impacts of children onto parents' time shares are one of my particular interests. In the next chapter of this thesis, I focus on the time-use changes that couples need to make at the extensive margin when they enter parenthood. The female labour supply literature has long documented the significant impact of motherhood onto women's career choice and developments (Blau and Kaln (1996), Weischselbaumer and Winter-Ebmer (2005), Lundborg, Plug, and Rasmussen (2017)). I am interested in not only which time-use category needs to make a cut for the extra time spent on the child but also how male partners may share the burden. This is particularly interesting given the recent event of COVID-19 pandemic which made many of us either stay at home or work from home for a long period of time.

1.2 Econometric framework

The seminal paper by Papke and Wooldridge (1996) was the first to accommodate fractional responses on its whole unit interval. Using a quasi-maximum likelihood approach, the authors allowed the fractional outcome variables to take any values on the $[0, 1]$. This popular method has since been applied for many cross-sectional studies and become increasingly popular thanks to its computational simplicity and intuitive appeal. However, it does not accommodate a panel setting and the extension is non-trivial.

The panel setting is one of the key feature of my estimators, beside non-linearity. So far the setting has been only for cross-sectional data where one observes N individuals at a given moment in time. As it turns to the situation where one has access to a panel data set, i.e. where the same individuals are observed repeatedly over time, the model can take this extra dimension of longitudinal data to tease out the individual fixed effects. This component plays a significant role in my model as it is likely to correlate with variables of interest and hence, causes estimation bias if ignored.

The extension of the fractional response model from a cross-section to a panel setting could take the path of either random effects (RE) or fixed effects (FE). There have been fractional outcome models using RE approach for different forms of non-linearity such as logit (Podivinsky and Stewart (2009)), probit (Papke and Wooldridge (2008)), and tobit (Loudermilk (2007), Elsas and Florysiak (2015)). It is noted that the RE approach does require more substantial distributional assumptions compared to the FE one. However, the trade-off with FE lies in the difficulty of dealing with the individual effects once there is a non-linear function such as logit involves. In the scope of the FE models, I either have to estimate individual intercepts α_i as part of the specification or find a way to eliminate them, most commonly by applying a transformation.

The first FE approach is to estimate the fixed effects. It has been taken by Hausman and Leonard (1997) proposing an updated quasi-maximum likelihood approach for estimating the fixed effects. However, it still suffers the incidental parameters problem (IPP). Neyman and Scott (1948) have defined the problem as one where the estimation of the parameters of interest may depend on the incidental ones. Hence, if the incidental parameters could not be estimated consistently, neither would the parameters of interest. In a panel data setting which has a small fixed number of time periods, T, and a relatively large number of observations, N, as the number of individual specific intercepts α_i increases at the exact same rate with N, I would not be able to get consistent estimates for α_i and more importantly, for coefficients of explanatory variables of interest.

The second FE approach is to eliminate the fixed effects. Although the elimination could be done fairly easy with linear setting by using the within transformation, there is no universal solution for non-linear models. The transformation technique is largely dependent on the functional form. Bonhomme (2012) proposed a systematic approach to construct moment restrictions on parameters of interest that are free from the individual fixed effects. However, it generally relied on likelihood models where extra assumptions are required.

For the logistic function, there has been a solution proposed by Chamberlain (1980), but only for binary logit data. The full extension to accommodate fractional data has not been resolved completely. In addition, the actual practical implementation was still left unclear. A recent work by Ramalho, Ramalho, and Coelho (2016) has been able to purge the individual effects, however, at a cost of values at one of the boundaries. Winkelmann and Xu (2019) had a unique approach of transforming the fractional data to discrete form and subsequently used the existing methods of count data. However, this method limits to univariate outcome and rational fractions only.

Consequently, motivated by this gap in the literature, my estimators offer a unique technique to solve this estimation issue. To the best of my knowledge, compared to the existing literature for univariate fractional model under panel setting, the proposed method is the first that could manage to have the following two characteristics at the same time: (i) respecting and handling the inherent bounds of $[0, 1]$ for the dependent variable simultaneously and (ii) directly eliminating the fixed effects while requiring minimal assumption of the conditional mean having a logistic form.

1.3 General caveats of the chosen data and approach

In terms of data, the longitudinal HILDA data set provides me the key to handle fixed effects. Although it comes with great facility to deal with personal heterogeneity, the data was self-recalled with no precise dairy outlay of when the activity happens. Due to its nature as a survey, the data collection is done every year in a fixed time frame. The respondents therefore only recall about an average amount of time they spent in a typical week over the whole year retrospectively. This could potentially have a large measurement error. With the use of fixed effects, potentially persistent bias caused by individuals overestimating or underestimating their time for particular categories is eliminated. Previous studies have stated that men appear to overestimate their time spent in housework more so than women (Baxter and Bittman (1995), Press and Townsley (1998), Kan and Pudney (2008)).

In addition, it also does not record any multitasking activities. While this could be picked up in a time diary, survey respondents may double count their time spent on both categories if they actually multi-tasked. For example, mothers could both do houseworks and keep an eye on her children. She may only need to occasionally attend her child should there be any help needed. Subsequently, there exists a limit in the quality of time use data and the collection method. It is understood that they tend to be very time-consuming and costly to satisfy all of the aspects. Given the focus of my study, the advantages of longitudinal features outweigh its compromises.

In terms of approach, as my choice of the functional form is the multinomial logit, it bears the inherent feature that discrimination among alternatives reduces to a series of pairwise comparisons. They are unaffected by the characteristics of alternatives other than the pair under consideration. This well-known limitation of the multinomial logit, usually termed Independence of Irrelevant Alternatives (IIA), implies zero correlation between the disturbances of the utilities associated with the various alternatives. In the fractional context, however, an analogous statement of independence between the errors of shares equations is not consistent with the unit-sum nature of the responses. For any pair of time shares, we may see that both shares may change proportionately in the event of a change in other options of time shares, or they may not. Therefore, the Independence of Irrelevant Alternatives is not a plausible assumption for the application of multinomial logit on fractional outcomes. Further discussion of IIA is presented in Chapter 2 with my I theoretical attempt at relaxing this assumption subsequently in Chapter 3. In sum, IIA is a caveat associated with my choice of multinomial logit as the selected functional form carries the implausible assumption when being adopted to fractional framework.

1.4 Preview of the chapters and key findings

In Chapter 2, after constructing the estimators for binary case, I further extend them to accommodate for multinomial fractional variables. The structure of the Generalised Method of Moments (GMM) condition allows the transition to multinomial setting proceed without difficulties. This extension now allows my model to handle data with a great versatility. I can apply one simple estimator for many types of logit-based variables, from univariate binary to multinomial fractional. In addition, it can concurrently handle individual effects without dropping any data points. This feature sets the method apart from previous ones which were only for a cross-sectional platform Mullahy (2015). The newly developed estimators provide a convenient and simple tool for a wide range of empirical research. It facilitates particularly panel studies which require multinomial shares having a non-trivial probability at the boundary values.

In my findings, women take a larger impact onto their time allocations when the couple enters parenthood. It shows that the estimate of the negative impact onto their work hours from having a newborn baby is much larger compared to that of their male partner. Although the gap tends to reduce over time and also as the child grows up, it still persists. I also find stronger impact of first born than any higher order births. Besides, couples with higher education, higher income, and especially, more progressive attitude towards gender's role appear to react stronger with the fathers allocating much more time to childcare than their peers in families which have a more traditional view.

There is a natural shortcoming associated with my choice of functional form for the conditional mean, in particular, the multinomial logit. As part of the construction, the standard multinomial logit would carry a strong assumption of Independence of Irrelevant Alternatives (IIA). Mixed logit is one of the most commonly adopted approaches to relax this condition. In Chapter 3, I therefore aim to obviate this shortcoming by extending my fractional models, in a similar manner with the discrete choice literature, to a mixed logit model, which hence allows my framework to be fully flexible.

I attempt to measure and estimate the distribution and density of the random coefficients for fractional responses with fixed effects. I document the performance of each method and explain possible reasons behind these findings. I find that none of the methods could precisely estimate the variation of the random coefficients. It however shows that the original method could identify the mean. In addition, I find that the simulated moment condition is very sensitive to the magnitude of disturbance, from the heterogeneity of the random coefficients and also from the conditional expectation error. My chosen functional form of the conditional mean could potentially lead to a reduction in tractability of the model.

In Chapter 4, I look at the correlation structure among couples' time allocations across work and home activities. I focus on different components driving the coordination in time allocations. I classify potential factors into three categories: unobservable individual preferences, observable events, and unknown time varying factors. While it is fairly straight forward to measure the contribution of observable factors, it is less clear among the unknowns: individual preferences and time-varying residuals. I hence disaggregate them by adapting the Poisson Fixed Effects model. Subsequently, I derive the contribution towards time allocation correlation structure from these three different aspects.

I find that individual preferences show strong correlations across many time shares. It signifies the similarity within each couple in terms of views towards gender's role and division of labour. I also document strong joint structure due to external events. Finally, the time varying unknowns do not appear to hold any correlation patterns. It subdues concerns towards their role in determining correlations among the time share outcomes.

Finally, Chapter 5 concludes my thesis by summarising the key contributions and findings of the main chapters.

In summary, my thesis contributes to the literature of fractional response models in a panel setting by proposing a quasi-differencing method which highlights the importance of individual heterogeneity. Across the three substantive chapters, I first introduce the construction of the estimators and how it may be applied for an empirical study of time uses where individual fixed effects play an important role. I then explore the possibility of extend the estimators to handle individual effects not only at the intercept but also at the slope coefficients, which comes hand in hand with the attempt to relax one of the key assumptions. Although the simulation results are unfavourable, I believe my second study has provides further insights for the literature of mixed logit and the limitations my proposed estimator may encounter when the true form of slope parameters are not constants. In my third and final study, I explore different assumptions around the error term and illustrate an interesting cross-correlation approach for the same topic of time uses. Overall, the thesis aims to provide a set of alternative and novel approaches for fractional response models and household time allocations literature.

Chapter 2

Couples' time use before and after the kids: A fractional panel fixed effects approach

2.1 Introduction

In the rich literature related to female labour participation, it has long been argued that being a mother has a significant negative impact on employment. In spite of a closing gender gap over the last decades, women in almost all developed countries still have less income than men, take longer for career development and are a minority in leadership positions Blau and Kaln (1996), Weischselbaumer and Winter-Ebmer (2005). Being a mother is a significant factor attributed to these disadvantages. For example, Herr (2015) recorded a relationship between delayed first birth and higher post-birth labour supply while Lundborg et al. (2017) found a "large and long lasting negative effect" of fertility on mothers' careers. The cost of children may involve career breaks, reduced hours of work, decreased life-time earnings, and at the same time, be accompanied with extra hours of child care and housework.

As the female labour supply literature mainly categorises market versus non-market hours, the non-paid activities remain as one single group. This study looks at the finer details of these non-paid works which could include childcare, house work and the remaining time. Concurrently with market hours, the study aims to measure the impact of parenthood on these multiple time uses simultaneously. However, it is noted that, to allow each partners time allocations to be explicitly dependent upon the decisions of their partner, it may require a substantive game-theoretic maximisation of a joint household utility function, which this study does not attempt to accomplish. It rather provides an improvement in the econometric approach which practitioners may use to include the partners' variables equivalently in the equation.

When entering motherhood, women's time reallocation may result from a joint household coordination rather than a decision made. In many cases, the outcome results from a joint household coordination with their partners. Ignoring this interdependent role of partnership will hence omit an important aspect of the joint structure. However, the existing literature of labour supply usually does not model the role of both partners concurrently. For many studies, the partner's information is often found on the right-hand-side of the regression as some covariates. In my approach, both partners time allocations on various different categories are modelled as outcome variables.

The theoretical frameworks for partners' joint time allocation have been developed as part of the household production function literature. While there are different approaches to study the way households operate, I focus on time use allocation and adapt the framework developed by Rogerson and Wallenius (2019) (RW) for this study. Although there have been other models such as Chiappori and Mazzocco (2017), I find that RW's framework has features most similar to my econometric settings. Not only do we both have multi-categories of time shares, my estimators also provide a direct interpretation for their labour supply elasticities. Their formulation is later shown to be represented as a function of my estimates.

Within the time shares literature, the use of information on how people spend their time has been increasingly popular. The development could be attributed to the improvement in availability of data. Hamermesh (2016) provided an excellent review of the current works in the topic and where it could be extended to. Time, as a resource, is inevitably increasingly scarce. As we may have continuing growth in other types of resources, the amount of time we each have could hardly increase. Depending on one's preferences and personal characteristics, he or she might allocate their time differently. Thus, the presence of individual heterogeneity is substantial and may play a significant part in the decision making process. My study will highlight this key matter and develop models around the focal point of individual fixed effects.

Despite the potentially important role of individual effects, the amount of time use studies which accommodate this feature has been very limited. Hamermesh (2016) pointed out that most of the literature was conducted in a cross-sectional setting (Argyrous and Rahman (2017), Bauer and Sonchak (2017)) which is not equipped to tease out the unobservable individual effects. This could lead to potential omitted variable bias if these unobservables correlate to the explanatory variable of interest.

This limitation mainly results from the lack of availability of time use data in a panel setting. Most collections of time-use diaries have only been repeated cross sections. This lack of individual longitudinal information causes validation issues for these studies and incentivised researchers to look for a solution: panel survey Craig and Siminski (2011), Foster and Stratton (2018). Although there are drawbacks due to larger measurement error by its self-reporting nature, the advantages of panel time use data outweigh, especially in this context where correlated individual effects are of high concerns. Taking advantage of the rich longitudinal data from the Household, Income and Labour Dynamics in Australia (HILDA) survey, this study contributes to a growing literature in time-use that implemented panel features for their studies.

To fill these gaps, I will need an estimator that could address individual fixed effects for a non-linear model which has multinomial share outcomes. To the best of my knowledge, there is no current econometric model that could satisfy these requirements concurrently. This motivated me to build an econometric model that is capable of accommodating these aspects. I focus on the case of panel data sets which have a large N and a small fixed T. This particular setting allows me to highlight the severity of the problem caused by genuine fixed effect approach where dummy variables are included to represent fixed effects (Greene (2002)).

In my findings, women take a larger impact onto their time allocations when the couple enters parenthood. It shows that estimate of the negative impact onto their work hours from having a newborn baby are much larger compared to that of their male partner. Although the gap tends to reduce over time and also as the child grows up, it still persists. I also find stronger impact of first born than any higher order births. Besides, couples with higher education, higher income, and especially, more progressive attitude towards gender's role appear to react stronger with the fathers allocate much more time shares than their peers in traditional families.

The remainder of this chapter is set as follow. Section 2.2 and 2.3 introduce the literature review and the theoretical econometric framework of my estimates followed by Section 2.4 which illustrates the Monte Carlo simulation comparing small sample performance of my estimators with others in a variety of settings. Section 2.5 explains the extension to the RW's household allocation model and presents the empirical application results before I conclude in the final Section 2.6.

2.2 Fractional responses

In many economic settings, the response variable (y) is often a proportion or a fraction being defined on the closed interval of $[0; 1]$, i.e. $0 \leq y \leq 1$. Examples of fractional responses are debt ratios, participation rates, market shares, pass rates, television ratings, and fraction of farming land Hausman and Leonard (1997), Brownstone and Train (1999), Elsas and Florysiak (2015). The bounded nature of these variables and, in many cases, the possibility of nontrivial probability mass at the boundaries raise some estimation and inference issues.

For many years, there have been three main approaches to model fractional response variables. Consider a random sample of $i = 1, \ldots, N$ individuals and let y be the fractional variable of interest, $0 \leq y \leq 1$, and $\mathbf{x}$ a vector of k covariates. Let β be the vector of parameters to be estimated.

The first of them, often used by many empirical researchers, is simply to ignore the bounded nature of y and assume a linear conditional mean model for y. However, this standard practice of using linear models to examine how a set of explanatory variables x influences a given proportional or fractional response variable y is not appropriate in general, since it does not guarantee that the predicted values of the dependent variable are restricted to the unit interval.

$$\mathbb{E}[y|\mathbf{x}] = \mathbf{x}\beta \tag{2.1}$$

This linear specification cannot ensure that the predicted values of y lie between 0 and 1 without severe constraints on the range of x or ad-hoc adjustments to fitted values outside the interval.

Aware of these problems, some empirical researchers opted for assuming the logistic relationship

$$\mathbb{E}[y|\mathbf{x}] = \frac{e^{\mathbf{x}\beta}}{1 + e^{\mathbf{x}\beta}} \tag{2.2}$$

which is indeed a natural choice for modelling proportions since it ensures that $0 < \mathbb{E}[y|\mathbf{x}] < 1$. However, instead of estimating the logistic equation directly, which would require some nonlinear technique, most authors prefer to estimate by least squares. The log-odds ratio model is defined by

$$\mathbb{E}[\log \frac{y}{1-y}|\mathbf{x}] = \mathbf{x}\beta \tag{2.3}$$

which corresponds to the linearisation of the equation that results from solving $y = e^{\mathbf{x}\beta}/1 + e^{\mathbf{x}\beta}$ with respect to $\mathbf{x}\beta$. This approach has a main drawback. The transformed dependent variable is not well defined for the boundary values 0 and 1 of y, requiring ad-hoc adjustments if such values are observed in the sample, such as adding an arbitrarily chosen small constant to all observations of y.

In addition, when there are many observations at the upper and lower limits of the response variable, it is relatively common to use Tobit models for data censored at one and zero. However, there is a conceptual problem with this approach. As some authors argue (e.g. Maddala (1991)), a Tobit model is appropriate to describe censored data in the interval $[0, 1]$ but its application to data which is defined only in that interval is not easy to justify: observations at the boundaries of a fractional variable are a natural consequence of individual choices and not of any type of censoring. Moreover, the Tobit model is stringent in terms of assumptions, requiring normality and homoskedasticity of the dependent variable, prior to censoring.

Given the limitations of these approaches, some alternatives that account for the bounded nature of the variable of interest have been proposed. Two main approaches for modelling fractional data without boundary observations have been proposed so far. The first only requires the correct specification of the nonlinear conditional expectation of the fractional response variable. The second alternative is a fully parametric approach, where a particular conditional distribution is assumed for the fractional dependent variable. I choose the first approach as it requires minimal assumption for the outcome variables. Within the nonlinear conditional mean approach, some of them can only be used when there are no observations at the boundaries, while others may also be employed when one or both the limits are observed with a positive probability. However, all of them have in common the utilisation of functional forms for the conditional mean of y that enforce the conceptual requirement that $\mathbb{E}[y|\mathbf{x}]$ is in the unit interval.

The simplest solution for dealing with fractional response variables only requires the assumption of a functional form for y that imposes the desired constraints on the conditional mean of the dependent variable:

$$\mathbb{E}[y|\mathbf{x}] = G(\mathbf{x}\beta) \tag{2.4}$$

where $G(\cdot)$ is a known nonlinear function satisfying $0 \leq G(\cdot) \leq 1$. This approach was first formally proposed by Papke and Wooldridge (1996), who suggested as possible specifications for $G(\cdot)$ any cumulative distribution function. A popular choice for $G(\cdot)$ is the logistic function which, however, instead of being first linearised as discussed above, must be directly estimated using nonlinear techniques.

The model defined may be consistently estimated by nonlinear least squares (NLS), as in Hermalin and Wallace (1994)'s empirical application, or, as suggested by Papke and Wooldridge (1996), by quasi maximum likelihood (QML). The latter authors proposed a particular QML method based on the Bernoulli log-likelihood function, which is given by

$$LL_i(\beta) = y_i \log[G(\mathbf{x_i}\beta)] + (1 - y_i)\log[1 - G(\mathbf{x_i}\beta)] \tag{2.5}$$

As the Bernoulli distribution is a member of the linear exponential family

(LEF), the QML estimator of β defined by

$$\hat{\beta} = \arg\max_{\beta} \sum_{i=1}^{N} LL_i(\beta) \tag{2.6}$$

is consistent and asymptotically normal, regardless of the true distribution of y conditional on $\mathbf{x}$, provided that $\mathbb{E}[y|\mathbf{x}]$ is correctly specified.

This popular method has since been applied for many cross-sectional studies and become increasingly popular thanks to its computational simplicity and intuitive appeal. The most common of those methods, where the mean function $G(\cdot)$ takes the logistic form, has since been applied in numerous empirical studies, including Hausman and Leonard (1997), Liu et al. (1999), and Wagner (2001). However, Papke and Wooldridge (1996) does not accommodate a panel setting and the extension is non-trivial.

Hausman and Leonard (1997) applied fractional logit to panel data on television ratings of National Basketball Association games to estimate the effects of superstars on telecast ratings. In using pooled QMLE with panel data, the only extra complication is in ensuring that the standard errors are robust to arbitrary serial correlation (in addition to misspecification of the conditional variance). But a more substantive issue arises with panel data and a non-linear response function: How to account for unobserved individual heterogeneity that is possibly correlated with the explanatory variables.

However, even as Hausman and Leonard (1997) included fixed effects in the model to account for unobserved heterogeneity, it still suffers the incidental parameters problem (IPP). The IPP problem is defined as one where the estimation of the parameters of interest may depend on the incidental ones. In the case where the incidental parameters could not be estimated consistently, neither would the parameters of interest Neyman and Scott (1948).

Wagner (2001) analyses a large panel data set of firms to explain the export-sales ratio as a function of firm size. Wagner explicitly included firm-specific intercepts in the fractional logit model, a strategy suggested by Hardin and Hilbe (2007) when one observes the entire population (as in Wagners case, because he observes all firms in an industry). Generally, while including dummies for each cross section observation allows unobserved heterogeneity to enter in a flexible way, it suffers from the IPP under random sampling when T (the number of time periods) is small and N (the number of cross-sectional observations) is large. In particular, with fixed T, the estimators of the fixed effects are inconsistent as $N \to \infty$, and this inconsistency transmits itself to the coefficients on the common slope parameters. As this study focuses on the case of panel data sets which have a large N and a small fixed T, it complements the existing work by Machado (2004) where the author works with cases where N is rather limited.

It is common that empirical researchers often have to deal with the problem of unobserved individual heterogeneity in their econometric models. When the model is linear in the parameters, this issue is easily dealt with. For example, if the omitted variables are uncorrelated with the variables included in the model, then unobservables may be simply ignored, and standard application of ordinary least squares (OLS) produces unbiased estimators of the parameters of interest. If, instead, unobserved and observed covariates are correlated, then, provided that a set of instruments is available for the endogenous regressors, instrumental variables based approaches, such as the generalized method of moments (GMM), may be applied.

In the framework of non-linear models, it is much more complicated to deal with unobserved heterogeneity and its consequences are in general more serious. Overall, authors that do incorporate the heterogeneity in the model in a sensible manner then typically choose one of the following strategies: (i) make strong distributional assumptions for the unobservables, which often generate poorly fitting models; (ii) work with linearised versions of the model of interest (e.g., log-transformed models for non-negative responses), which, typically, cannot be directly applied in cases where boundary values of y are observed with non-zero probability; or (iii) find a unique and innovative way, depending on each non-linear functional form, to transform the conditional mean so that the key parameters of interest could be estimated without depending on the individual fixed effects.

In this paper, I focus on the approach of the third strategy. Let y_{it} denote the fractional response variable, defined on the interval $[0;1]$, to be explained for individual i, $i = 1, \ldots, N$, at time t, $t = 1, \ldots, T$, and let $\mathbf{x_{it}}$ denote a k-vector of explanatory variables. A direct extension of the conditional mean function in a cross sectional setting in Papke and Wooldridge (1996) to a fixed-effects panel data setting yields

$$\mathbb{E}[y_{it}|\mathbf{x_{it}}] = G(\mathbf{x_{it}}\beta + \alpha_i) \tag{2.7}$$

where α_i denotes time-invariant unobserved heterogeneity. While under correct specification of $G(\cdot)$ it is straightforward to obtain consistent estimators for β in the cross-sectional model (Papke and Wooldridge (1996)), the same does not happen with the panel data model due to the presence of unobservables in the argument of the $G(\cdot)$ function. All the alternatives suggested, namely estimating explicitly the parameters α_i (Hausman and

Leonard (1997)), assuming a distribution for α_i (Loudermilk (2007); Papke and Wooldridge (2008); Wooldridge (2010);Elsas and Florysiak (2015)) or using a linearised model based on log-odds transformations, have several drawbacks.

The focus of Ramalho et al. (2016) is on the estimation of structural parameters when the response variable is defined on the interval $[0; 1)$. Given the difficulty of obtaining consistent estimators for β under the panel setting, the authors propose new estimators for cases where the data generating process (DGP) of y_{it} may be described in a slightly different way. In particular, the regression model proposed in the next section requires the DGP to be given by

$$y_{it} = G(\exp(\mathbf{x_{it}}\beta + \alpha_i + \epsilon_{it})) \tag{2.8}$$

where ϵ_{it} denotes time-varying unobserved heterogeneity and $G(\cdot)$ is assumed to have a functional form $G(x) = \frac{x}{1+x}$.

Let $H(\cdot) = G(\cdot)^{-1}$. Then, (2.8) can be expressed as an exponential model with a transformed dependent variable:

$$H(y_{it}) = \exp(\mathbf{x_{it}}\beta + \alpha_i + \epsilon_{it}) \tag{2.9}$$

where $H(y_{it}) = \frac{y_{it}}{1-y_{it}}$

A quasi first-difference transformation is obtained:

$$\frac{H(y_{it})}{\exp(\mathbf{x_{it}}\beta)} - \frac{H(y_{i,t-1})}{\exp(\mathbf{x_{i,t-1}}\beta)} = \xi_{it} \tag{2.10}$$

The model parameter β is then estimated in the GMM framework, utilising the moment condition of:

$$\mathbb{E}\left[\frac{H(y_{it})}{\exp(\mathbf{x_{it}}\beta)} - \frac{H(y_{i,t-1})}{\exp(\mathbf{x_{i,t-1}}\beta)}|\mathbf{x_{it}}\right] = 0 \tag{2.11}$$

Another innovative approach was proposed by Winkelmann and Xu (2019). Instead of finding a new estimator for the fractional outcomes, the authors have transformed the fractions into a set of indicator variables. Therefore, they can now apply existing methods for discrete models. The key underlying idea to make this approach work is to keep the conditional expectation function stay the same during the transformation and hence utilising existing standard estimators for discrete model.

The binomial logit fixed effects estimator can be implemented using any off-the-shelf statistical software with a conditional logit routine, since the binomial distribution arises as the sum of K independent Bernoulli trials. Therefore, two estimators are equivalent: one based on a binomial log-likelihood function and the other based on a Bernoulli log-likelihood for an expanded dataset.

For the expanded dataset, one simply generates a sequence of K copies for each i, keeping the regressors unchanged, where y_{it} is replaced by a sequence of 0/1 indicator variables d_{ijt} in arbitrary order such that

$$\sum_{j=1}^{M} d_{ijt} = My_{it} \tag{2.12}$$

It follows that d_{ijt} and y_{it} have the same conditional mean:

$$\mathbb{E}\left[y_{it}|\mathbf{x_{it}}\right] = \mathbb{E}\left[\frac{\sum_{j=1}^{M} d_{ijt}}{M}|\mathbf{x_{it}}\right] = \mathbb{E}\left[d_{ijt}|\mathbf{x_{it}}\right] \tag{2.13}$$

The logit (Bernoulli) log-likelihood function of the expanded dataset is given by

$$\begin{aligned}\log L &= \sum_{i=1}^{N}\sum_{j=1}^{M} d_{ijt}\log(\Lambda_{it}) + (1-d_{ijt})\log(1-\Lambda_{it}) \\ &= \sum_{i=1}^{N} Y_{it}\log(\Lambda_{it}) + (M-Y_{it})\log(1-\Lambda_{it})\end{aligned} \tag{2.14}$$

This log-likelihood function is proportional to the binomial log-likelihood as well as to the Bernoulli quasi-log-likelihood (Papke and Wooldridge (1996), replacing Y_{it} by y_{it} an $(M-Y_{it})$ by $(1-y_{it})$, and the three maximum likelihood estimators are therefore identical.

My estimator lies within this third approach where it offers a unique technique to eliminate the time-invariant unobservables away from the conditional mean. To the best of my knowledge, compared to the existing literature for univariate fractional model under panel setting, the proposed method is the first that could manage to have the following two characteristics at the same time: (i) respecting and handling the inherent bounds of $[0, 1]$ for the dependent variable simultaneously and (ii) directly eliminating the fixed effects while requiring minimal assumption of the conditional mean having a logistic form.

I further extend my estimator to accommodate multinomial fractional variables. I can apply one simple estimator for many types of logit-based variables, from univariate binary to multinomial fractional. In addition, it can concurrently handle individual effects without dropping any data points. The newly developed estimators provide a convenient and simple tool for a wide range of empirical researches. It facilitates panel studies which require multinomial shares having a non-trivial probability at the boundary values.

I demonstrate the ability of the proposed fractional logit estimator to perform well under a range of data generating processes (DGPs) where other estimators struggle to deliver useful estimates, as well as to investigate the finite sample performance of the estimator against a benchmark in cases where an efficient alternative estimator is available. Throughout my simulation results, the newly developed estimator has outperformed the other benchmarks. In the baseline DGP, the pooled estimator has suffered a large bias of 44 per cent over the true value while the Poisson estimator increased its biasedness as the mass of data points at boundary 1 increases.

2.3 Econometric framework

In this section, I propose a class of estimators within the framework of generalised method of moments (GMM) for panel data. The novelty of the contribution lies in exploring the unique features of the logistic fractions to obtain moment conditions that are free from individual fixed effects. As highlighted above in Section 2.1, these GMM moment conditions can accommodate values at both zero and one bounds into .

2.3.1 Quasi-differenced fractional logit estimator

The panel dataset comprises of dependent variable, y_{it}, and a vector of independent variables, $\mathbf{x}'_{it}$, where $i = 1, ..., N$ indicates individual i and $t = 1, ..., T$ indicates time t. y_{it} is a fraction defined in the closed interval of $[0, 1]$. It may take on the values of 0 or 1. Binary outcome is a special case of this set-up.

The analysis begins with the nonlinear specification of the conditional expectation function (CEF) for y_{it}:

$$\mathbb{E}[y_{it}|\alpha_i, \mathbf{x}_{it}, \beta] = \frac{\exp(\alpha_i + \mathbf{x}'_{it}\beta)}{1 + \exp(\alpha_i + \mathbf{x}'_{it}\beta)} \equiv \Lambda_{it} \tag{2.15}$$

where β is a vector of parameters of interest and α_i is the fixed effects for each individual.

I first focus on the case of $T = 2$, where it is not only the most tractable but it also is the scenario where the IPP bias is most severe: $y_{i1} = \Lambda_{i1} + \epsilon_{i1}$ and $y_{i2} = \Lambda_{i2} + \epsilon_{i2}$ where ϵ_{it} are mean-zero CEF errors. The IPP problem is defined as one where the estimation of the parameters of interest may depend on the incidental ones. In the case where the incidental parameters could not be estimated consistently due to a small fixed T, neither would the parameters of interest Neyman and Scott (1948). As the number of incidental paramaters increase at the exact same rate with N.

Instead of transforming y_{it}, I look at the pairwise products of y_{it} and

$1 - y_{is}$, with $\{s, t\} \in \{1, 2\}$ in this case of $T = 2$:

$$\begin{aligned} y_{i1}(1 - y_{i2}) &= (\Lambda_{i1} + \epsilon_{i1})(1 - (\Lambda_{i2} + \epsilon_{i2})) \\ &= (\Lambda_{i1} + \epsilon_{i1})((1 - \Lambda_{i2}) - \epsilon_{i2})) \\ &= \Lambda_{i1}(1 - \Lambda_{i2}) + \epsilon_{i1}(1 - \Lambda_{i2}) - \epsilon_{i2}\Lambda_{i1} - \epsilon_{i1}\epsilon_{i2} \end{aligned} \tag{2.16}$$

$$\begin{aligned} y_{i2}(1 - y_{i1}) &= (\Lambda_{i2} + \epsilon_{i2})(1 - (\Lambda_{i1} + \epsilon_{i1})) \\ &= (\Lambda_{i2} + \epsilon_{i2})((1 - \Lambda_{i1}) - \epsilon_{i1})) \\ &= \Lambda_{i2}(1 - \Lambda_{i1}) + \epsilon_{i2}(1 - \Lambda_{i1}) - \epsilon_{i1}\Lambda_{i2} - \epsilon_{i1}\epsilon_{i2} \end{aligned} \tag{2.17}$$

Observe that

$$\frac{\Lambda_{i1}(1 - \Lambda_{i2})}{\Lambda_{i2}(1 - \Lambda_{i1})} = \exp((\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\beta)$$

is free from α_i, from (2.16) and (2.17),

$$\begin{aligned} & y_{i1}(1 - y_{i2}) - \exp((\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\beta) y_{i2}(1 - y_{i1}) \\ =& y_{i1}(1 - y_{i2}) - \frac{\Lambda_{i1}(1 - \Lambda_{i2})}{\Lambda_{i2}(1 - \Lambda_{i1})} y_{i2}(1 - y_{i1}) \\ =& [\Lambda_{i1}(1 - \Lambda_{i2}) + \epsilon_{i1}(1 - \Lambda_{i2}) - \epsilon_{i2}\Lambda_{i1} - \epsilon_{i1}\epsilon_{i2}] \\ & - \frac{\Lambda_{i1}(1 - \Lambda_{i2})}{\Lambda_{i2}(1 - \Lambda_{i1})} [\Lambda_{i2}(1 - \Lambda_{i1}) + \epsilon_{i2}(1 - \Lambda_{i1}) - \epsilon_{i1}\Lambda_{i2} - \epsilon_{i1}\epsilon_{i2}] \\ =& \epsilon_{i1}(1 - \Lambda_{i2}) - \epsilon_{i2}\Lambda_{i1} - \epsilon_{i1}\epsilon_{i2} \\ & - \frac{\Lambda_{i1}(1 - \Lambda_{i2})}{\Lambda_{i2}(1 - \Lambda_{i1})} [\epsilon_{i2}(1 - \Lambda_{i1}) - \epsilon_{i1}\Lambda_{i2} - \epsilon_{i1}\epsilon_{i2}] \end{aligned} \tag{2.18}$$

Taking expectation gives us the moment condition of:

$$\mathbb{E}[y_{i1}(1 - y_{i2}) - \exp((\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\beta) y_{i2}(1 - y_{i1}) | \mathbf{x}_{i1}, \mathbf{x}_{i2}, \alpha_i] = 0 \tag{2.19}$$

This is achieved under standard assumptions of no auto-correlation among the CEF errors and exogeneity between them and the explanatory variables (Wooldridge (2019)).

Let $u_i = y_{i1}(1-y_{i2}) - \exp((\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\beta)y_{i2}(1-y_{i1})$ and $z_i = \mathbf{x}_{i1} - \mathbf{x}_{i2}$ As the expression is free of individual fixed effects, I could rewrite the conditional moment as:

$$\mathbb{E}[u_i|z_i] = 0 \tag{2.20}$$

Hence, we have the corresponding the unconditional moment:

$$\mathbb{E}[u_i z_i] = 0 \tag{2.21}$$

The conditional sample analogue of (2.21) could be illustrated below

$$\frac{1}{N}\sum_{i=1}^{N}[u_i z_i] = 0 \tag{2.22}$$

My GMM estimator would hence be the $\hat{\beta}$ that solves this sample analogue moment condition (2.22).

For the case of a fixed $T > 2$, there would be $\frac{1}{2}T(T-1)$ potential pairwise moment conditions. I choose to use consecutive pairs of t and $t+1$. This follows tha same pratice which was implemented for ArellanoBond estimator.

2.3.2 Extension to multinomial fractional

I further extend my GMM estimators to accommodate panel setting which has multinomial fractional dependent variable. By adopting the multinomial logit for the conditional mean, I can re-utilise the same quasi-differencing method for the multinomial case. It carries the transformation in a similar manner to a univariate version.

The panel dataset now is $(y_{it}^{[k]}, \mathbf{x}'_{it})$, $i = 1, ..., N$, $t = 1, ..., T$, $k = 1, ..., K+1$ where $y_{it}^{[k]}$ denotes the k^{th} fractional response variable of interest, i.e. each lies in the interval $[0, 1]$ and adds up to 1 $(\sum_{k=1}^{K+1} y_{it}^{[k]} = 1)$. The last $K + 1^{th}$ category is omitted leaving the model K dependent variables. Similarly, any of these could take the boundary values of 0 or 1 with non-trivial masses.

The multinomial version utilises the multinomial-logit based conditional expectation function of:

$$\mathbb{E}[y_{it}^{[k]}|\mathbf{x}'_{it}, \alpha_i^{[k]}] = \frac{\exp(\alpha_i^{[k]} + \mathbf{x}'_{it}\beta^{[k]})}{1 + \sum_{l=1}^{K} \exp(\alpha_i^{[l]} + \mathbf{x}'_{it}\beta^{[l]})} \equiv \Lambda_{it}^{[k]} \qquad \text{where } k = 1, \ldots, K$$

and

$$\mathbb{E}[y_{it}^{[K+1]}|\mathbf{x}'_{it}, \alpha_i^{[K+1]}] = \frac{1}{1 + \sum_{l=1}^{K} \exp(\alpha_i^{[l]} + \mathbf{x}'_{it}\beta^{[l]})}$$

where $\beta^{[k]}$ is a vector of parameters of interest and $\alpha_i^{[k]}$ is the fixed effects for each individual and each time use type.

Through similar transformation as in the univariate case, I have the analogue of the moment condition for the multinomial setting:

$$\frac{1}{N} \sum_{i=1}^{N} [u_i^{[k]} z_i^{[k]}] = 0 \tag{2.23}$$

The multinomial logit conditional mean bears the inherent feature that discrimination among alternatives reduces to a series of pairwise comparisons which are unaffected by the characteristics of alternatives other than the pair under consideration. This well-known limitation of the standard multinomial logit is referred to as the independence of irrelevant alternatives (IIA). In the fractional context, it is an implausible assumption by construction. The changes in one time share may impact on the rest of the fractional outcomes.

2.4 Monte Carlo experiments

The aim of the Monte Carlo experiments is to demonstrate the ability of the proposed quasi-differenced logit estimator (QD) compared with other estimators under a range of different data generating processes (DGPs). As there exists many different potential settings for the DGPs, the following exercise could only cover a certain number of them. The experiments also investigate the finite sample performance of QD estimator against other benchmarks in cases where an efficient alternative is available.

2.4.1 Univariate fractional dependent variable

The first class of DGPs I consider are for a single fractional dependent variable $y_{it} \in [0, 1]$. The DGPs comprise normally distributed fixed effects α_i and a single regressor x_{it}:

$$\begin{aligned} \alpha_i &\sim N(0, \sqrt{1/2}), \\ x_{it} &= \alpha_i + \xi_{it}, \text{ where } \xi_{it} \sim N(0, \sqrt{1/2}), \quad i = 1, \ldots, N, \ t = 1, \ldots, T. \end{aligned}$$

This specification implies a correlation between fixed effects and regressor which is independent of the number of time periods T and which amounts to about 70 per cent ($\rho = 0.5/\sqrt{0.5} \approx 0.707$). The dependent variable is generated as the share of successes of a binomial random variable with success probability Λ_{it},

$$y_{it} = w_{it}/c, \text{ where } w_{it} \sim B(c, \Lambda_{it}).$$

Table 2.1. MC Simulation: Univariate DGP ($\beta_1 = 1$)

	$\beta_0 = -2$			$\beta_0 = 0$			$\beta_0 = 2$		
	Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
I. *Baseline: N=10,000, T=2, c=10,* $\rho \neq 0$									
Pool	1.438	0.030	0.439	1.430	0.025	0.431	1.438	0.034	0.439
Pool with FE	1.138	0.022	0.421	1.130	0.028	0.411	1.138	0.044	0.429
Pois	0.974	0.072	0.076	0.801	0.056	0.207	0.554	0.042	0.448
QD	1.000	0.050	0.050	1.004	0.058	0.058	0.997	0.069	0.069
II. *T=8, rest equal to baseline*									
Pool	1.441	0.018	0.441	1.429	0.017	0.430	1.440	0.017	0.440
Pool with FE	1.137	0.026	0.422	1.131	0.029	0.431	1.139	0.042	0.431
Pois	0.956	0.030	0.053	0.737	0.020	0.264	0.479	0.016	0.521
QD	1.001	0.025	0.025	0.998	0.020	0.020	0.999	0.024	0.024
III. *c=30, rest equal to baseline*									
Pool	1.440	0.023	0.440	1.429	0.019	0.429	1.438	0.025	0.439
Pool with FE	1.135	0.021	0.415	1.122	0.025	0.405	1.135	0.041	0.429
Pois	1.035	0.075	0.083	0.943	0.060	0.083	0.721	0.043	0.283
QD	1.001	0.031	0.031	1.000	0.032	0.032	0.998	0.038	0.038
IV. α_i*=0, rest equal to baseline*									
Pool	1.002	0.020	0.021	0.998	0.018	0.018	1.000	0.024	0.024
Pool with FE	1.138	0.022	0.421	1.130	0.028	0.411	1.138	0.044	0.429
Pois	1.051	0.057	0.076	0.926	0.044	0.086	0.484	0.039	0.517
QD	1.005	0.057	0.057	0.996	0.050	0.051	0.990	0.066	0.067
V. *c=1 (binary response variable), rest equal to baseline*									
Pool	1.442	0.086	0.450	1.426	0.065	0.431	1.448	0.075	0.454
Pool with FE	1.132	0.021	0.420	1.129	0.028	0.413	1.120	0.044	0.421
Clogit	1.011	0.170	0.170	1.002	0.146	0.147	1.015	0.150	0.151
QD	1.057	0.242	0.248	1.017	0.217	0.217	1.054	0.242	0.248

Notes:Cell entries contain mean (Mean), standard deviation (SD) and root mean squared error (RMSE) for $\beta_1 = 1$ estimated over 100 replications for the three estimators Pool (pooled logit MLE), Pois (Poisson PLE) and QD (Quasi-differenced PLE). The parameter β_0 is the intercept. In all panels, the number of individuals is $N = 10,000$. T indicates the number of time periods, c the number of binomial trials over which the fractional response variable is calculated, and ρ the correlation between α_i and the regressor x_{it}. Thus, for $\alpha_i = 0$, ρ=0.

Thus, the success probability is the mean of the fractional variable y_{it}. We specify

$$\Lambda_{it} = \Lambda(\beta_0 + \beta_1 x_{it} + \alpha_i),$$

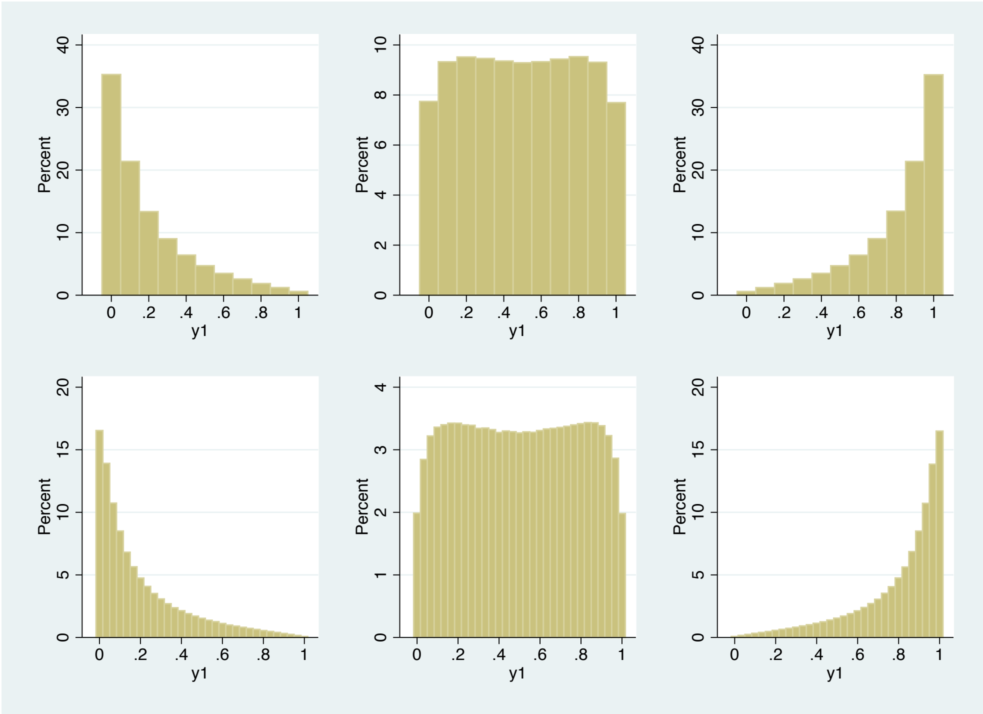
and my main interest lies in the estimation of β_1, which I set to $\beta_1 = 1$ throughout all DGPs. In my baseline DGP, I set the number of individuals $N = 10,000$ reflecting the size of the actual empirical study and the number of time periods $T = 2$ where the issue of IPP is most severe.

The parameter c determines the support points of my dependent variable. For the case $c = 1$, the DGP is a default binary logit. As c is increased, y_{it} has support points inside the $[0, 1]$. In my baseline case I set $c = 10$, which implies $y_{it} \in \{0, 0.1, 0.2, \ldots, 1\}$. The parameter β_0 controls the share of bound or limit observations. I specify $\beta_0 = -2, 0, 2$. The first and last case give substantial shares of observations with $y_{it} = 1$ and $y_{it} = 0$, respectively, while $\beta_0 = 0$ results in a moderate percentage of both 0 and 1 observations.

Figure 2.1 shows three different distributions for the dependent variable. The focus of the design is to show three different cases of fractions I may observe in practice. It may evenly distributed as the second case, where $\beta_0 = 0$ or skewed to one side as $\beta_0 = -2$ or $\beta_0 = 2$. I then embed these scenarios into different DGPs where four estimators would be compared against each other.

I consider five estimators for this model. First, the pooled logit which ignores the fixed effects (Pool). This is the Papke-Wooldridge Logit PL cross-section estimator applied to panel data without any modifications. While estimates can be obtained for any c, it is only consistent in the case $\alpha_i = 0$ for all i. Second, I include fixed effects as dummy variables in the list of

Figure 2.1. MC Simulation–Univariate DGP: Distribution of y_{it} for baseline DGP and DGP with $c = 30$



Notes: The top panel shows the distribution of y_{it} when $c = 10$ versus that of y_{it} in the bottom panel when $c = 30$, the figure also compares across 3 different values of $\beta_0 \in -2, 0, 2$ which control the shapes of the distribution, and thickness of the boundary values.

regressors (Greene (2002)). This highlights the issues of IPP in cases where I have a large N and a small fixed T. Third, I consider the Poisson FE PL estimator with dependent variable $\tilde{y}_{it} = y_{it}/(1 - y_{it})$ (Pois). This is the Hausman-Hall-Grilliches (Hausman, Hall, and Griliches (1984)) Poisson fixed effects estimator applied to $\tilde{y}_{it}$. It is closely related to the Ramalho-Ramalho-Murteira (Ramalho, Ramalho, and Murteira (2011)) family of estimators, but for the case where the conditional expectation function is logistic. It is consistent for any DGP with $y_{it} \in [0, 1)$; or, if y_{it} is appropriately redefined, for any DGP with $y_{it} \in (0, 1]$, but never for the entire interval $[0, 1]$. The

"Pois" estimator is based on the conditional expectation function.

$$E(\tilde{y}_{it}|x_{it},\alpha_i) = E\left(\frac{y_{it}}{1-y_{it}}\right) = \exp(\beta_0 + \beta_1 x_{it} + \alpha_i),$$

and can therefore be estimated by any fixed effects exponential mean model in principle. Taking the logarithm of $\tilde{y}_{it}$ would make it possible to estimate the model by OLS but it would imply further restricting the possible support of y_{it} to $(0,1)$. Because it cannot handle both bounds, the "Pois" estimator cannot be applied to the binary logit case: after losing observations at one bound, no variation is left in the dependent variable. For this case, however, I consider a fourth estimator, Chamberlain's conditional logit estimator (Clogit). However, this estimator cannot handle fractional responses, so it is only applicable to DGPs with $c = 1$. Finally, the fifth and last estimator I consider is the novel quasi-differenced logit PL estimator. This estimator has the advantage that it can handle both bounds and is consistent for any c.

Throughout Table 2.1, the QD estimator has outperformed the other benchmarks. In the baseline DGP, the pooled estimator has suffered a large bias of 44 per cent over the true value of $\beta_1 = 1$ while the Poisson estimator increased its biasedness as the mass of data points at boundary 1 increases. We observe the same behaviour for the next two cases where $T = 8$ and $c = 30$. In the fourth case where individual fixed effects are assumed away, the pooled estimator slightly outperformed QD thanks to its higher efficiency. Lastly, I show that in the case of binary response, my estimator can achieve a reasonably good performance, although it is less accurate than Clogit which was designed for this specific setting. These results show that my estimator is the only method that could achieve reasonably desired results across all cases. However, it is not always necessarily the best in terms of efficiency due to its designated construction.

2.4.2 Multinomial fractional response variable

The second class of DGPs I consider are for a multinomial fractional response variable with three outcomes $\boldsymbol{y}_{it} = (y_{it}^{(1)}, y_{it}^{(2)}, y_{it}^{(3)})'$, with $y_{it}^{(1)} + y_{it}^{(2)} + y_{it}^{(3)} = 1$ and $y_{it}^{(k)} \in [0,1]\ \forall k = 1,2,3$.

We specify

$$\Lambda_{it}^{(1)} = \frac{\exp(z_{it}^{(1)})}{1+\exp(z_{it}^{(1)})+\exp(z_{it}^{(2)})}, \qquad \Lambda_{it}^{(2)} = \frac{\exp(z_{it}^{(2)})}{1+\exp(z_{it}^{(1)})+\exp(z_{it}^{(2)})},$$

where

$$z_{it}^{(1)} = \alpha^{(1)} + \beta_0^{(1)} + \beta_1^{(1)} x_{it}, \qquad z_{it}^{(2)} = \alpha^{(2)} + \beta_0^{(2)} + \beta_1^{(2)} x_{it},$$

as the respective means of $y_{it}^{(1)}$ and $y_{it}^{(2)}$. As before I obtain the dependent variables as shares of successes from binomial random variables,

$$y_{it}^{(1)} = w_{it}^{(1)}/c, \quad y_{it}^{(2)} = w_{it}^{(2)}/c, \quad \text{where} \quad w_{it}^{(1)} \sim B(c, \Lambda_{it}^{(1)}), \quad w_{it}^{(2)} \sim B(c, \Lambda_{it}^{(2)}).$$

The fixed effects are generated as

$$\alpha_i^{(1)} = \alpha_i, \quad \alpha_i^{(2)} = -1.5\alpha_i,$$

where α_i is taken from the univariate DGP. The regressor x_{it} is also generated in the same way as in the univariate DGP. Here, I specify $\beta_1^{(1)} = 1$ and

$\beta_1^{(2)} = -1$ as my main objects of interest. As before, the constants determine the share of bound observations. I keep $\beta_0^{(1)} = \beta_0^{(2)}$ throughout, and set values of -2,-1 and 0.

Across both of the DGPs, the QD estimator outperforms the pooled estimator. As there are no other equivalent estimator in the scope of multinomial setting, I keep the baseline DGP and the binary response variable with the benchmark estimator - pooled. Table 2.2 shows a consistent deviation of pooled estimates away from the true values for both β_1 and β_2. It reaffirms the importance of handling fixed effects and proves a consistent performance of my estimator when being extended to multinomial setting.

Table 2.2. MC Simulation: Multinomial DGP ($\beta_1 = 1$ and $\beta_2 = -1$)

		$\beta_0 = -2$			$\beta_0 = -1$			$\beta_0 = 0$		
		Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
I. *Baseline: N=10,000, T=2, c=10,* $\rho \neq 0$										
Pool	$\hat{\beta}_1$	1.398	0.031	0.400	1.355	0.029	0.356	1.308	0.040	0.311
	$\hat{\beta}_2$	-1.562	0.039	0.563	-1.507	0.038	0.508	-1.451	0.043	0.453
Pool	$\hat{\beta}_1$	1.198	0.021	0.390	1.315	0.029	0.356	1.308	0.040	0.311
with FE	$\hat{\beta}_2$	-1.362	0.039	0.563	-1.507	0.038	0.508	-1.451	0.043	0.453
QD	$\hat{\beta}_1$	0.990	0.074	0.075	0.991	0.071	0.071	0.993	0.084	0.084
	$\hat{\beta}_2$	-1.004	0.080	0.080	-1.002	0.071	0.071	-0.997	0.084	0.084
V. *c=1 (binary response variable), rest equal to baseline*										
Pool	$\hat{\beta}_1$	1.397	0.093	0.408	1.353	0.077	0.362	1.313	0.115	0.333
	$\hat{\beta}_2$	-1.563	0.090	0.570	-1.510	0.085	0.517	-1.461	0.109	0.473
Pool	$\hat{\beta}_1$	1.198	0.021	0.390	1.315	0.029	0.356	1.308	0.040	0.311
with FE	$\hat{\beta}_2$	-1.361	0.038	0.563	-1.507	0.038	0.508	-1.451	0.043	0.453
QD	$\hat{\beta}_1$	1.018	0.260	0.261	1.055	0.309	0.314	1.098	0.411	0.422
	$\hat{\beta}_2$	-1.070	0.327	0.335	-1.097	0.400	0.411	-1.054	0.354	0.358

Notes:Cell entries contain mean (Mean), standard deviation (SD) and root mean squared error (RMSE) for $\beta_1 = 1$ and $\beta_2 = -1$ estimated over 100 replications for the two estimators Pool (pooled logit MLE) and QD (Quasi-differenced PLE). The parameter β_0 is the intercept. In all panels, the number of individuals is $N = 1,000$. T indicates the number of time periods, c the number of binomial trials over which the fractional response variable is calculated, and ρ the correlation between α_i and the regressor x_{it}. Thus, for $\alpha_i = 0$, ρ=0.

2.5 Application

Through the previous section, the performance of my QD estimator has been proven across many different simulation settings. In this section, the estimators are then used to measure the impact of children on couples' time use. The exercise aims to estimate the simultaneous effect of having a baby onto the couples' time shares of paid work, child care, house work and leisure.

Our data is extracted from the Household, Income and Labour Dynamics in Australia (HILDA) Survey where the sample selects working aged mixed-gender couples. With these features, the application provides a multinomial fractional response setting which helps illustrating how my estimators could be used in an empirical works.

Started in 2001, the annual HILDA Survey provides a unique insight into the dynamics of Australian labour market. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). Beside its focus on labour market outcomes, it covers a wide range of topics such as education, health, fertility, retirement, family relationship, etc.

The sample of analysis consists of all observations of respondents aged 15 to 65 years with complete mixed-gender couples profile for at least two waves. I started with 91,280 observations spanning over 15 years. It is noted that this is an unbalanced panel as existing respondents leaving and new ones entering the survey throughout this period of time. Overall, there were 14,907 unique individuals identified. First, I removed 14,034 observations that are out of standard working age, who are more than 65 or less than 15 years of age. Second, I dropped 2,518 due to partner's missing data. I also excluded 52 observations that did not report the same number of children under the same household. Finally, I used couples that appeared at least twice in my data. Hence, a further 2,258 data points dropped which results in a final number of 25,519 couples. We could observe that while male's time shares for paid work and leisure are significantly larger than those of their female partners, they spend less time on child care and housework.

Table 2.3. Descriptive Statistics

	Male	Female	Difference
Time share - work	0.256	0.145	0.111***
	(0.121)	(0.119)	(0.001)
Time share - child care	0.043	0.093	-0.050***
	(0.061)	(0.132)	(0.001)
Time share - housework	0.059	0.138	-0.079***
	(0.049)	(0.088)	(0.001)
Time share - leisure	0.642	0.623	0.019***
	(0.142)	(0.165)	(0.001)
Age	44.30	42.07	2.23***
	(11.12)	(10.87)	(0.097)
1 if highest education level is college	0.309	0.353	-0.044***
	(0.462)	(0.478)	(0.004)
1 if being employed	0.879	0.721	0.157***
	(0.327)	(0.448)	(0.003)
Log of hourly wage	3.009	2.569	0.440***
	(1.033)	(1.285)	(0.010)
Annual non-labour income	16108	10762	5346***
	(55209)	(37355)	(417)
N	25519	25519	

Notes: The first two columns show the mean and standard deviation in parentheses. The last column indicates the difference in mean between male and female with standard error in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Another observation is that female partners earn less hourly wage, having lower non-labour income although possessing higher education level.

Before discussing the results, I will visit the model of household time allocation derived from Rogerson and Wallenius (2019) with a fitted adjustment for my own study. In this subsection, I show how my results could be interpreted in a theoretical framework of household time allocation model.

2.5.1 Model of Household Time Allocation

In the original model of Rogerson and Wallenius (2019), the authors considered 3 different types of time shares for each member j ($j \in \{male, female\}$) of a mixed-gender coupled family i at time t: market work m_{it}^j, house-work h_{it}^j and leisure l_{it}^j with $m_{it}^j + h_{it}^j + l_{it}^j = 1$. For the purpose of my study, I further detach child care hours share from the leisure, notated as ch_{it}^j. Therefore, under my setting, I instead have a simplex of 4: $m_{it}^j, ch_{it}^j, h_{it}^j$ and l_{it}^j where $m_{it}^j + ch_{it}^j + h_{it}^j + l_{it}^j = 1$ for each member j.

The maximisation problem for the household i is characterised as follow:

$$\max u^i(c_{it}, l_{it}) \text{ s.t. } g_{it} + a_{it} = w_{imt}m_{imt} + w_{ift}m_{ift} + (1 + r_t)a_{it-1} \tag{2.24}$$

where the consumption $c_{it} = [\alpha_i g_{it}^{1-1/\eta} + \beta_i ch_{it}^{1-1/\eta} + (1 - \alpha_i - \beta_i)h_{it}^{1-1/\eta}]^{\frac{\eta}{\eta-1}}$ combines g_{it} being the household expenditure while child care hours ch_{it} and housework hours h_{it} being shared between the male and female partners: $ch_{it} = [A_i^m(ch_{it}^m)^{1-1/\rho} + A_i^f(ch_{it}^f)^{1-1/\rho}]^{\frac{\rho}{\rho-1}}$ $h_{it} = [B_i^m(h_{it}^m)^{1-1/\sigma} + B_{it}^f(h_{it}^f)^{1-1/\sigma}]^{\frac{\sigma}{\sigma-1}}$. Finally, I have household leisure hour share l_{it} also expressed in the same sharing manner: $l_{it} = C_i^m \frac{(l_{it}^m)^{1-1/\delta}}{1-1/\delta} + C_{if}^f \frac{(l_{it}^f)^{1-1/\delta}}{1-1/\delta}$.

From the maximisation problem, I derive the first order conditions with respect to h_{it}^m and h_{it}^f which give us:

$$\begin{aligned} \frac{\partial u^i(c_{it}, l_i t)}{\partial h_{it}^m} &= u_1^i(c_{it}, l_{it})c_{it}^{\frac{1}{\eta}} h_{it}^{\frac{1}{\rho}-\frac{1}{\eta}} B_i^m (h_{it}^m)^{\frac{-1}{\rho}} - u_2^i(c_{it}, l_{it})C_i^m(1 - m_{it}^m - ch_{it}^m - h_{it}^m)^{\frac{-1}{\delta}} = 0 \\ \frac{\partial u^i(c_{it}, l_i t)}{\partial h_{it}^f} &= u_1^i(c_{it}, l_{it})c_{it}^{\frac{1}{\eta}} h_{it}^{\frac{1}{\rho}-\frac{1}{\eta}} B_i^f (h_{it}^f)^{\frac{-1}{\rho}} - u_2^i(c_{it}, l_{it})C_i^f(1 - m_{it}^f - ch_{it}^f - h_{it}^f)^{\frac{-1}{\delta}} = 0 \end{aligned} \tag{2.25}$$

Combining these two conditions in (2.25), I simplify it to:

$$\frac{B_i^m}{B_i^f}\left[\frac{h_{it}^m}{h_{it}^f}\right]^{\frac{-1}{\rho}} = \frac{C_i^m}{C_i^f}\left[\frac{l_{it}^m}{l_{it}^f}\right]^{\frac{-1}{\delta}} \tag{2.26}$$

Similarly for child care hour shares ch_{it}^m and ch_{it}^f, I derive:

$$\frac{A_i^m}{A_i^f}\left[\frac{ch_{it}^m}{ch_{it}^f}\right]^{\frac{-1}{\sigma}} = \frac{C_i^m}{C_i^f}\left[\frac{l_{it}^m}{l_{it}^f}\right]^{\frac{-1}{\delta}} \tag{2.27}$$

The two conditions (2.26) and (2.27) pin down the equality in the marginal rate of substitution between leisure of the two members with that of their housework shares and also that of their child care shares. Although it has certain structure, these equations impose no restriction on the value of key elasticity parameters of ρ, σ and δ if I only have cross-sectional data. Hence, by having access to panel data, I could further rearrange (2.26) and (2.27) into:

$$\begin{aligned}\left[\frac{h_{it}^m}{h_{it}^f}\right]^{\frac{-1}{\rho}}\left[\frac{l_{it}^m}{l_{it}^f}\right]^{\frac{1}{\delta}} &= \frac{C_i^m}{C_i^f}\frac{B_i^f}{B_i^m} \\ \left[\frac{ch_{it}^m}{ch_{it}^f}\right]^{\frac{-1}{\sigma}}\left[\frac{l_{it}^m}{l_{it}^f}\right]^{\frac{1}{\delta}} &= \frac{C_i^m}{C_i^f}\frac{A_i^f}{A_i^m}\end{aligned} \tag{2.28}$$

As the values on the right hand side of (2.28) are assumed to be time invariant, taking natural logs and first differencing gives:

$$\Delta\log\frac{l_{it}^m}{l_{it}^f} = \frac{\delta}{\rho}\times\Delta\log\frac{h_{it}^m}{h_{it}^f} = \frac{\delta}{\sigma}\times\Delta\log\frac{ch_{it}^m}{ch_{it}^f} \tag{2.29}$$

The two member household model resulted in a simple set of conditions which only involves time shares and key labour supply elasticity parameters. This hence allows a direct interpretation of our estimates as it involves the same natural logarithm forms and settings. As seen in equation 2.29, the relation is limited to the relative term between the male and female partners for each type of time allocations. The empirical setting allows further flexibility as I could simultaneously find the elasticity across couple and across different time use categories.

We take female leisure l_{it}^f as the reference category, hence, having the other seven groups $m_{it}^m, ch_{it}^m, h_{it}^m, l_{it}^m, m_{it}^f, ch_{it}^f, h_{it}^f$ numbering $y_{it}^{[k]}$ with k running from 1 to 7. In general, I have my estimator of interest $\beta^{[k]}$ being interpreted as the percentage change of $y_{it}^{[k]}$ relative to female leisure given a change in X_{it}:

$$\frac{\Delta \log \frac{y_{it}^{[k]}}{l_{it}^f}}{\Delta \mathbf{x}_{it}'} = \beta^{[k]} \text{ and } \frac{\Delta \log \frac{y_{it}^{[k]}}{y_{it}^{[l]}}}{\Delta \mathbf{x}_{it}'} = \beta^{[k]} - \beta^{[l]} \text{ where } k, l = 1, \ldots, 7 \tag{2.30}$$

Using the notation of (2.30), I rewrite the household time allocation optimisation condition of (2.29) in terms of $\beta^{[k]}$ as:

$$\begin{aligned} \frac{\delta}{\rho} &= \frac{\beta^{[4]}}{\beta^{[3]} - \beta^{[7]}} \\ \frac{\delta}{\sigma} &= \frac{\beta^{[4]}}{\beta^{[2]} - \beta^{[6]}} \end{aligned} \tag{2.31}$$

I can thus calculate these key labour elasticity parameters of RW model directly from my estimators.

$$\frac{\hat{\delta}}{\rho} = \frac{\hat{\beta}^{[4]}}{\hat{\beta}^{[3]} - \hat{\beta}^{[7]}} \text{ and } \frac{\hat{\delta}}{\sigma} = \frac{\hat{\beta}^{[4]}}{\hat{\beta}^{[2]} - \hat{\beta}^{[6]}} \tag{2.32}$$

Through equation 2.32, the estimates which subsequently produced in the empirical section provides the necessary inputs for the calculates presented in RW model. With respect to the interpretation, although they differ with one being labour elasticities whereas the other representing relative partial effects, they all link together through the above set of equations.

2.5.2 Results

I first present some key results from the baseline setting before exploring different configurations to the outcomes of interest. I also compute the average marginal effects which are comparable to linear regression studies.

Baseline

In the baseline specification, I have the main variables of interest based on the number of dependent children in the family by different age group. Although the focus is on the youngest group where children are between 0 to 4 years of age, the larger age group also reveal useful insights of how parents coordinate their time allocations as their children grow. Table 2.4 shows a consistent reduction of impacts onto mothers' hours of work and hours of childcare as her children gets older. One interesting point is noted that the impact seems to converge between the parents when their children reach the teenage phase.

Beside the key variables of children, other labour participation related variables also provide interesting insights. The first one is the effect of log wage ratio of female versus male partner on the time shares. From Table 2.4, it is shown to be large and significant. Relative to female's leisure, her work share relatively increases significantly (11.6%) for every unit change in the log wage ratio. At the same time, she reduces her child care and house work share (-3.2% and -0.8% respectively) while her partner works less and does more house work (-1.4% and 2.1% respectively). These figures have shown the importance of labour opportunity cost as a factor of the household optimisation problem. With men traditionally being the bread earners, they have had a significantly higher labour market value compare to their counterparts. A positive shift in a relative term of income for women could see a significant reallocation across the two members of the family. Next is the annual non-labour household income which reflects the reduction of work hours once families receive other sources of income apart from their wages. I see that this income decreases the work hour shares proportionately in a similar way for both members (about -3%), however, none of the other time allocation categories responds to this variable. Third, all of the estimates for the effect of the local unemployment rate are insignificant and close to zero in magnitude. It suggests that the macroeconomic condition does not explain patterns of time allocations in this sample. Lastly, I took control of possible trend effect by including the time dummies and observe a monotonic inter-temporal trend across all the time shares.

Table 2.4. Baseline results

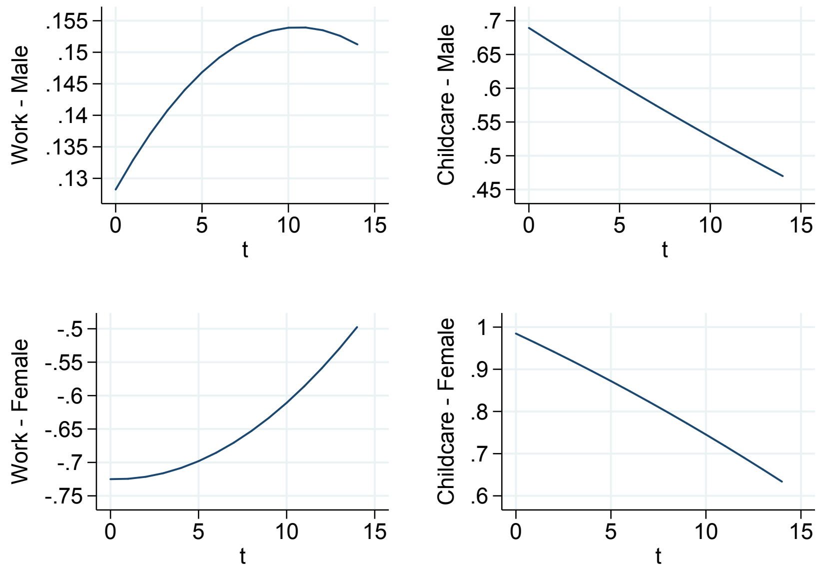
	Male				Female		
	Work	Child	House	Leisure	Work	Child	House
No. of kids (0-4 years old)	0.143*** (0.010)	0.586*** (0.029)	0.215*** (0.021)	0.093*** (0.008)	-0.631*** (0.037)	0.800*** (0.028)	0.339*** (0.019)
No. of kids (5-9 years old)	0.080*** (0.010)	0.457*** (0.031)	0.129*** (0.023)	0.055*** (0.008)	-0.289*** (0.026)	0.602*** (0.031)	0.209*** (0.020)
No. of kids (10-14 years old)	0.055*** (0.009)	0.366*** (0.034)	0.071*** (0.021)	0.035*** (0.006)	-0.140*** (0.020)	0.467*** (0.034)	0.150*** (0.019)
No. of kids (15-24 years old)	0.036*** (0.007)	0.239*** (0.034)	0.055** (0.017)	0.015** (0.005)	-0.024 (0.014)	0.294*** (0.035)	0.068*** (0.015)
Log wage ratio	-0.014*** (0.003)	0.007 (0.007)	0.021*** (0.004)	0.008*** (0.001)	0.116*** (0.007)	-0.032*** (0.007)	-0.008* (0.004)
Annual non-labour income (hh)	-0.033*** (0.006)	-0.002 (0.019)	-0.005 (0.012)	0.005 (0.004)	-0.030** (0.010)	-0.008 (0.020)	0.004 (0.011)
Local unempl. rate	-0.003 (0.004)	-0.008 (0.014)	0.006 (0.009)	-0.002 (0.002)	0.009 (0.007)	-0.025 (0.015)	-0.007 (0.008)
Time dummies	Yes		No. of couples				5396
No. of observations	25519		J statistic				0.000

Standard errors in parentheses

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

Notes: The results show the corresponding estimated coefficients of each covariate for the 7 fractional outcomes. The omitted fraction is female's leisure. These coefficients have similar interpretation with those from the multinomial logit regression as they share similar conditional mean. The difference is that these are fractions, not probabilities of realised discrete outcomes.

Relative to female leisure, I observe in Table 2.4 a significant reduction in paid work for both partners and particularly, a large drop in childcare for women. The ratio between her child care share relative to leisure is halved over the course of 15 years. This, however, could be attributed by technology advancement in time-saving household appliances and availability of childcare outsourcing. It is nonetheless a positive progression for women as they could now free up more unpaid work time to reallocate them into



Notes: The figures show temporal changes in the impacts of children aged between 0 and 4 onto relative work share and child care share for males and females over the last 15 years.

Figure 2.2. Impact of 0 - 4 year old child over time

paid work or leisure.

In Figure 2.2, I examine how the effect of young children on parental time use changed over the past 15 years. Figure 2.2 shows the quadratic temporal effect from the youngest age group 0 - 4 year old children on their parents' work and childcare time uses. On the left panel, both of the parents appear to work more compared to 15 years ago in the event of entering parenthood. The effect of children onto fathers' work hour shares increased all the way up to year 2012 before it levelled off afterwards. For mothers, the negative impact onto their work share reduced significantly by about 30% (from -0.75 in 2002 down to -0.5 in 2016). The right panel shows a monotonic relative reduction in the impact of a newborn baby onto their parents' child care time over the last 15 years in Australia. Altogether, this figure documented a reallocation of time use from child care to paid work over the last decade and a half in Australian households. Apart from generational differences in personal preference towards rearing and taking care of children that are decades away from each other, the figure documented a progressive trend in Australia promoting and helping couples to have children. Though this could partly attributed by a more available child care services also heavily subsidised by the government, the trend shows that being parents are, at the very minimum, a bit less burden than ten to fifteen years ago.

40-hour work week

In the next specification, I redefine my fractional response with a closer focus on a so-called marketable block of working hours per week. This also reflects a natural prioritised order of reactions in terms of time allocations. With work hours being the most difficult to rearrange and negotiate, I impose an order of flexibility across different time shares to represent this empirical factor. Hence, the time uses will now have the following order of flexibility, from the least to the most: paid work, child care, house work and lastly, leisure.

When individuals juggle their time, there would be components that are harder than easier to change. It could be the employment contract that one has already agreed to. This phenomena leads to a natural prioritised order of reactions: first I adjust my work hours, then later changes other unpaid work around this new employment arrangements including childcare, housework and finally leisure. Second, this reflects the idea that not all of one's unpaid hours are marketable, the labour market only demands up to a certain threshold. In this case, I choose a natural cut off at 40 hours not only due to it being the standard full-time work, including mid-session breaks, one could expect to supply to the market each week on average but also it is observable from my data set too.

In this set-up, I give each partner a block of $A = 40$ prime unit of marketable hours each week. The 40 hour-week reflects the standard full-time equivalent market production hours available for an individual on a typical weekly basis including breaks in the middle of the working day. This way, I redefined my fractional response as follow: There are four components that one could select to fill up this amount of allocated time: paid work production (W_{it}^j), child care (CH_{it}^j), home production (H_{it}^j) and leisure (L_{it}^j), however, with a presumed order of priority.

Due to the nature of fixed hours and binding contracts, paid work hours are prioritised to be filled up first. If $W_{it} \geq A$, then the shares are $(w_{it}^j, ch_{it}^j, h_{it}^j, l_{it}^j) = (1, 0, 0, 0)$. If $W_{it} < A$, I then fill up the rest of the allocated time with CH_{it}, H_{it} and subsequently L_{it}.

Therefore, if $W_{it} + CH_{it} \geq A$, then

$$(w_{it}, ch_{it}^j, h_{it}, l_{it}) = \left(\frac{W_{it}}{A}, 1 - \frac{W_{it}}{A}, 0, 0\right)$$

If $W_{it} + CH_{it} + H_{it} \geq A$, then

$$(w_{it}, ch_{it}, h_{it}, l_{it}) = \left(\frac{W_{it}}{A}, \frac{CH_{it}}{A}, 1 - \frac{W_{it} + CH_{it}}{A}, 0\right)$$

Finally, if $W_{it} + CH_{it} + H_{it} < A$, then

$$(w_{it}, ch_{it}, h_{it}, l_{it}) = \left(\frac{W_{it}}{A}, \frac{CH_{it}}{A}, \frac{H_{it}}{A}, 1 - \frac{W_{it} + CH_{it} + H_{it}}{A}\right)$$

.

Table 2.5. Time use: 40 Hours

	Male				Female		
	Work	Child	House	Leisure	Work	Child	House
No. of kids	0.805***	1.006***	0.403	0.434	-0.340	1.539***	0.577***
(0-4 years old)	(0.099)	(0.247)	(0.241)	(0.232)	(0.251)	(0.151)	(0.125)
No. of kids	0.596***	0.710*	0.065	0.207	-0.199	1.183***	0.450***
(5-9 years old)	(0.109)	(0.278)	(0.236)	(0.194)	(0.218)	(0.192)	(0.129)
No. of kids	0.341***	0.551*	0.281	0.285	-0.172	0.891***	0.376***
(10-14 years old)	(0.092)	(0.252)	(0.159)	(0.161)	(0.194)	(0.176)	(0.111)
No. of kids	0.175**	0.356	0.284	0.016	0.055	0.513**	0.140
(15-24 years old)	(0.061)	(0.278)	(0.145)	(0.128)	(0.149)	(0.187)	(0.074)
log wage ratio	0.042*	0.086	0.070***	0.084***	0.420***	-0.015	0.035*
	(0.017)	(0.066)	(0.017)	(0.015)	(0.055)	(0.043)	(0.014)
annual non-labour	-0.130**	0.478	-0.013	0.147	-0.320**	0.183	-0.024
income (hh)	(0.046)	(0.275)	(0.093)	(0.076)	(0.110)	(0.147)	(0.054)
local unemployment rate	-0.004	-0.035	0.021	-0.008	0.118	-0.032	-0.032
	(0.032)	(0.145)	(0.044)	(0.034)	(0.070)	(0.098)	(0.034)
Time dummies	Yes		No. of couples				5396
No. of observations	25519		J statistic				0.000

Standard errors in parentheses

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

Notes: The results show the corresponding estimated coefficients of each covariate for the newly constructed 7 fractional outcomes. The omitted fraction is the same - female's leisure. However, the fractions were built with an order of priority: work, child care, house work and finally, leisure.

Table 2.5 reports larger impacts onto fathers' work hours and both parents' childcare hours in the event of having an additional child. It is also noticed that different to the baseline result in Table 2.4, the effects of a larger child are much smaller with most of them are statistically insignificant. Overall, by taking into consideration a priority order across time allocations, I find most of the effects of having children falls onto work hours and childcare.

These results reflect the large change in young parents' time allocations once they enter their parenthood are in the juggling of work hours and child care responsibility. Inside the standard “marketable” lot of 40 hours, there have been adjustments in the event of having children that are far more stronger than those of the whole available 168 hours per week. The strong impacts illustrate that the struggle that young parents face when having a young child surrounds the arrangement of work and childcare hours.

In addition, it is interesting to observe a much quicker drop in the estimates of impacts when the children get older. Compared to the baseline case, they have much smaller effects onto their parents’ time allocations. It reflects the child’s independence once they reach the larger aged group and spend more time on their own or outside of the house. Parents at that point would no longer play a hands-on role in taking care of their child and certainly spending less of their time together with their children.

Birth order In this regression specification, I look at heterogeneous effects in birth order. In Table 2.6, I separate the number of dependent children in the two smaller aged groups by birth order: first, second and higher parity. I observe a significant portion of the total effects which is attributed only by the first-born rather than any of the higher ordered births.

Table 2.6. First and second births

	Male				Female		
	Work	Child	House	Leisure	Work	Child	House
First kid (0-4 years old)	0.295*** (0.019)	1.793*** (0.069)	0.358*** (0.040)	0.175*** (0.015)	-0.903*** (0.068)	2.201*** (0.062)	0.708*** (0.036)
First kid (5-9 years old)	0.184*** (0.026)	1.378*** (0.074)	0.220*** (0.054)	0.113*** (0.021)	-0.641*** (0.078)	1.733*** (0.077)	0.486*** (0.049)
Second kid (0-4 years old)	0.059*** (0.018)	0.075 (0.043)	0.143*** (0.039)	0.068*** (0.015)	-0.478*** (0.062)	0.319*** (0.042)	0.272*** (0.034)
Second kid (5-9 years old)	0.058* (0.028)	0.106 (0.067)	0.001 (0.057)	0.051** (0.019)	-0.245*** (0.067)	0.301*** (0.072)	0.222*** (0.049)
No. of higher parity kids (0-4 years old)	0.059** (0.022)	0.201*** (0.045)	0.096* (0.044)	0.025 (0.018)	-0.287*** (0.070)	0.295*** (0.052)	0.149*** (0.040)
No. of higher parity kids (5-9 years old)	0.048* (0.024)	0.230*** (0.060)	0.034 (0.047)	0.016 (0.018)	-0.133* (0.060)	0.409*** (0.071)	0.092 (0.047)
No. of kids (10-14 years old)	0.069* (0.027)	0.442*** (0.081)	0.109* (0.049)	0.043* (0.019)	-0.086 (0.067)	0.859*** (0.089)	0.166** (0.057)
No. of kids (15-24 years old)	-0.029 (0.025)	0.191 (0.102)	0.026 (0.066)	0.017 (0.018)	0.112* (0.051)	0.441*** (0.131)	-0.013 (0.068)
log wage ratio	-0.006 (0.004)	0.013 (0.010)	0.013 (0.008)	0.003 (0.003)	0.229*** (0.020)	-0.031*** (0.009)	-0.020** (0.007)
annual non-labour income (hh)	-0.011 (0.012)	0.002 (0.027)	-0.006 (0.022)	0.009 (0.008)	0.004 (0.026)	0.013 (0.030)	0.019 (0.023)
local unemployment rate	0.006 (0.008)	-0.015 (0.022)	0.010 (0.018)	-0.008 (0.006)	0.016 (0.019)	-0.026 (0.023)	-0.008 (0.016)
Time dummies	Yes		No. of individuals				1234
No. of observations	7029		J statistic				0.000

Standard errors in parentheses

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

Notes: The results show the corresponding estimated coefficients of each covariate for the 7 fractional outcomes. The omitted fraction is female's leisure. This time, I have changed the list of covariates so I can differentiate between the impacts of the first birth, the second birth and any later births onto parents' time uses.

The first childs birth is for most parents a profound experience carrying the potential to change life orientations and values. Hence, parents' time allocations are expected to change significantly for the first birth. As one could expect, for subsequent births, it may not have as high as an impact onto parents' time allocations. My result confirms that, i.e. the changes that you make at the extensive margin when you first become parents are much larger then the subsequent children. It seems to be the case that parents, especially mothers endure the greatest impact with their very first child. The entrance to parenthood appears to be the most challenging phase which requires more supports and resources then any other subsequent birth.

Partial effects

The general limitations of the fixed effects approach and non-linear functional form include complicated partial effects. First, due to having a fixed effects model, not only do I eliminate unwanted unobservable fixed effects but also all interesting time-invariant factors. In this context, I dropped interesting variables such as education, household income or gender's role attitude. Second, as I employ a multinomial logistic form instead of a linear function, I no longer have β directly interpreted as partial effects. The average partial effects are then calculated subsequently following the same process for multinomial logit.

$$\frac{\partial \hat{y}_{it}^{[k]}}{\partial x_{it}} = \hat{\beta}^{[k]} \hat{\Lambda}_{it}^{[k]} (1 - \hat{\Lambda}_{it}^{[k]}) \tag{2.33}$$

In Table 2.7, I present the average partial effects (APEs) for each type of household that has low versus high level of education (tertiary education versus secondary or lower), household income (top versus bottom quartile) or traditional versus progressive view towards gender's role (using a range of behavioural questions surrounding the role of women in taking responsibilities of housework and child care).

With APEs, I could also have a full overview of partial impacts of explanatory variables onto each of the time allocations. Through these figures, it is noted that only mothers' paid work and leisure shares are negatively affected by the children while all the other time allocations in the household increase. In terms of heterogeneous effects, I found significant larger APEs in fathers' time allocations when he was in higher educated, higher household income or more progressive attitude towards gender's role families. In particular, out of these three dimensions, progressive attitude seems to be the most important aspect as husbands in these households have their APEs twice as large for the impact on their time used for housework and childcare compared to those in more traditional families.

Table 2.7. Partial effects of having 0 - 4 year old child

	Male				Female			
	Work	Childcare	Housework	Leisure	Work	Childcare	Housework	Leisure
Couples' education								
Low	0.0064	0.0088	0.0042	0.0028	-0.0418	0.0262	0.0156	-0.0221
High	0.0105	0.0138	0.0034	0.0043	-0.0705	0.0460	0.0205	-0.0280
Household income								
Low	0.0059	0.0105	0.0056	-0.0030	-0.0313	0.0276	0.0158	-0.0310
High	0.0117	0.0123	0.0031	0.0059	-0.0726	0.0393	0.0205	-0.0202
Attitude towards gendser's role								
Traditional	0.0083	0.0073	0.0028	0.0041	-0.0375	0.0213	0.0168	-0.0232
Progressive	0.0081	0.0158	0.0052	0.0029	-0.0675	0.0417	0.0201	-0.0262

Notes: In this exercise, I divided my sample using their education, income and gender's role attitude. The results show the corresponding estimated average partial effects (APEs) of having a young child aged between 0 to 4 for all of the 8 fractional outcomes. They were calculated in the same way with multinomial logit.

As a joint household decision making process, women would benefit greatly from having a more progressive male partners towards her decisions between career commitment and child raising responsibility.

Ignoring individual heterogeneity

Table 2.8. Time use of households: Pooled Fractional Multinomial Logit

	Male				Female		
	Work	Child	House	Leisure	Work	Child	House
No. of kids	0.292***	1.060***	0.246***	0.080***	-0.170***	1.153***	0.369***
(0-4 years old)	(0.005)	(0.011)	(0.008)	(0.003)	(0.011)	(0.011)	(0.007)
No. of kids	0.151***	0.517***	0.127***	0.014***	-0.015	0.494***	0.220***
(5-9 years old)	(0.005)	(0.012)	(0.008)	(0.003)	(0.008)	(0.012)	(0.007)
No. of kids	0.145***	0.434***	0.110***	-0.001	0.040***	0.379***	0.172***
(10-14 years old)	(0.005)	(0.012)	(0.008)	(0.002)	(0.007)	(0.012)	(0.007)
No. of kids	0.133***	0.140***	0.067***	0.001	0.080***	0.036*	0.172***
(15-24 years old)	(0.005)	(0.013)	(0.009)	(0.002)	(0.008)	(0.014)	(0.007)
log wage ratio	-0.033***	0.038***	0.068***	0.032***	0.236***	-0.047***	-0.053***
	(0.002)	(0.006)	(0.004)	(0.001)	(0.004)	(0.006)	(0.003)
annual non-labour	-0.341***	-0.232***	-0.031**	0.034***	-0.479***	-0.235***	0.048***
income (hh)	(0.006)	(0.016)	(0.009)	(0.003)	(0.010)	(0.016)	(0.008)
local unemployment	-0.037***	-0.005	-0.009	0.004*	-0.033***	-0.038***	0.002
rate	(0.004)	(0.010)	(0.006)	(0.002)	(0.006)	(0.010)	(0.005)
Time dummies	Yes			J statistic			0.000
No. of observations	36209						

Standard errors in parentheses

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

Notes: The results were provided by the pooled approach, which ignores the longitudinal setting of the data set. I would like to show the gap between these results with the baseline results above. The differences indicate the important role of accounting for individual fixed effects.

In Table 2.8, I compare my baseline results with a pooled fractional multinomial logit regression. I wish to highlight the importance of handling fixed effects. I find that the results are very different to those in the baseline result of Table 2.4. For example, the negative impact of children onto women's work hour shares is only for the youngest aged group and then disappears for all the older child groups. The difference shows that as I are treating the whole panel data set as a cross section, ignoring the unobservable heterogeneous fixed effects, it could therefore produce omitted variable bias.

2.6 Conclusion

This chapter makes both methodological and empirical contributions. It is one of a few studies in the time allocation and female labour supply literature that feature a panel setting while allowing multiple categories of time shares as outcome variables. In order to suit this purpose, I proposed a simple logistic estimator which is the first in its class that can (i) handle boundary values at $y_{it} = 0$ and $y_{it} = 1$ without dropping any observations, (ii) eliminate fixed effects, and (iii) be used for multinomial outcome.

In sum, I found that the effects were much larger and persistent for women's time use, in particular, their work and child care time shares. I also found a stronger impact when imposing priorities over the time allocation categories. Next, the impact of the first born child is shown to be higher than any subsequent births. Finally, on average, couples with higher education, higher income, and especially, more progressive gender attitude, react stronger to parenthood. Male partners in progressive households especially involve far more than their peers in traditional families.

The caveats of IIA assumption for standard multinomial logit is carried onto my multinomial model. It motivates a further extension to mixed logit in the following Chapter 3 where I adopt mixed logit with random coefficients for my fractional outcome estimators. I present different approaches in attempts to measure the heterogeneity at the coefficient level.

Chapter 3

A mixed logit model for fractional responses

3.1 Introduction

In the last chapter, I developed a class of estimators that could be applied to a wide range of multinomial logistic conditional mean fractional outcomes. It is applicable in the context of all discrete or continuous, binary or multinomial, cross-section or longitudinal dependent variables. The combination of versatility and simplicity helps make the estimators available to a wide range of different data settings.

As previously introduced in Chapter 2, the boundedness of these fractional variables has created estimation issues due to its non-linearity. Although the linear model is straightforward and interpretable, it ignores the nature of bounded outcomes. For certain values of covariates, it provides out-of-range predictions for these fractional outcomes, and these values are unrealistic for interpretation. It is therefore preferable to specify a non-linear functional form for these types of data. In addition, the longitudinal setting demands specific methods to handle individual fixed effects. The common approaches are the random effects and fixed effects models. My proposed estimator offers a unique technique using the fixed effects pathway, under which it can accommodate more features of the required data than any other existing method.

I then further extended my estimator to accommodate multinomial fractional variables. With the structure of my logit conditional mean function and the quasi-differencing method, it allows the transition to the multinomial setting to proceed with ease. The extension now allows my model to handle data with a greater number of subcategories of three or more. This largely expands the types of data that the model could handle. I observe that fractional data sets frequently are comprised of more than two groups. For example, household time allocations usually have several types of categories such as work hours, child care, house work or leisure.

I can now apply one simple estimator for many types of logit-based variables, from univariate binary to multinomial fractional. This sets us apart further from the existing papers which were only for a single fraction. However, my model now relies on a condition - the independence from irrelevant alternatives (IIA). This hence motivates me to extend the model further to relax this assumption, concurrently providing another contribution in terms of modelling heterogeneity.

First, there is a natural shortcoming associated with my choice of multinomial logit functional form. When extending to multinomial fractional outcomes, I chose to use it for the conditional mean. As part of the construction, the multinomial logit carries a strong assumption of Independence of Irrelevant Alternatives (IIA). In this framework, IIA is standard and often relaxed by allowing the vector of coefficient of interest β to be i specific, i.e. β_i for each individual where $i = 1, \ldots, N$. The random coefficients will be mixed in a predetermined distribution. The model is referred to as "mixed logit" and is one of the most common solutions used to relax the IIA.

The term "mixed" reflects the choice probability specified as a mixture of logits over a specified mixing distribution Brownstone and Train (1999), McFadden and Train (2000). The development was motivated from the shortcomings of the multinomial logistic which the mixed logit model fully generalises by allowing for random taste variation, unrestricted substitution patterns, and correlation in unobserved individual-specific factors over time. All of these features were obviated in the standard model. As I base my quasi-differencing panel estimator for fractional response upon the same multinomial logistic conditional expectation function, these limitations remain accordingly. In this chapter, similarly to the discrete choice literature, I extend my fractional models to a mixed fractional logit where it overcomes the IIA.

Another one of the solutions is nested logit. To generate more realistic substitution patterns, nested logit models group those alternatives with common features into the same nests that exhibit greater substitution effects (Train, McFadden, and Ben-Akiva (1987)). The nested logit still assumes that within each nest the IIA assumption holds and relaxes IIA across different nests. Despite its ability to partially relax IIA, the nested logit suffers from significant shortcomings of sensitive parameter estimation due to different nesting structures (Kling and Thomson (1996)).

Mixed logit provides a full featured model for the discrete choice framework (Revelt and Train (1998)). Mixed logit generalises the conditional logit by introducing unobserved preference heterogeneity through the coefficients. This is accomplished by assuming a probability density function for β_i, $f(\beta_i|\theta)$ where θ is a vector of parameters. Conditional on $\beta_i^{[k]}$, the probability of selecting alternative k in the mixed logit is:

$$\Pr[y_{it} = k|x_{it}, \alpha_i, \beta_i^{[k]}] = \frac{\exp(x_{it}\beta_i^{[k]} + \alpha_i)}{\sum_{m=1}^{K} \exp(x_{it}\beta_i^{[m]} + \alpha_i)} \tag{3.1}$$

The mixed logit model allows the coefficient of interest to vary across individuals. As a result, instead of having a single estimate of β for all observations, it now specifies β_i to be individual specific where these parameters follow a determined distribution. In the discrete framework, the extension allows agents to now have variation in tastes, and these individual preferences would be assumed to follow a mixing distribution. The mixed choice probabilities are then built as an integral of the logistic functions from each individual over the pre-specified distribution.

This leads to the second contribution of the extension. Individual heterogeneity has only been modelled at the intercept level as the fixed effects α_i (with $i = 1, \ldots, N$) in the previous chapter. They were considered as one of the key components correlated with household decision making on time allocations. Intuitively, they reflect time-invariant preferences of each individual in their choice of allocating their time spent in daily activities with their partner. I introduced a novel method to handle the α_i in the setting of fractional panel data. However, the estimators did not consider individual effects further than the intercepts. In another word, it assumed the coefficients, or the key parameters β, to be constant across observations, implying that individuals would react in the same manner with each other. This rather strict assumption motivates me to further explore the possibility of having individual heterogeneity at the coefficient level. The extension to mixed logit now allows β_i vary across i, and presumed to follow a normal distribution with mean β and standard deviation σ_β. In the subsequent steps, I use an adaptive version of my proposed method onto this new setting.

To the best of my knowledge, this is the first paper attempts to measure and estimate the distribution and density of β_i for fractional responses with fixed effects. I document the performance of each method and explain possible reasons behind these findings. I find that none of the methods could precisely estimate the variation of the random coefficients β_i. It however shows that the original method could identify the mean of β_i. In addition, I find that the simulated moment condition is very sensitive to the magnitude of disturbance, from σ_β and especially from the CEF error noise σ_ϵ and my chosen functional form could potentially make it even more difficult to trace down the shape of the mixing distribution.

3.2 Review of the fractional logit and mixed logit

In this section, I will briefly summarise the estimators proposed in the previous chapter and then explain the added feature of mixed logit. Following the usual multinomial framework for discrete literature, I now extend my multinomial set-up to include random parameters along the lines of a "fractional mixed logit".

3.2.1 Fractional logit estimator

In the previous chapter, I started with a simple panel dataset of univariate fractional outcome $(y_{it}, \mathbf{x}'_{it}),\ i = 1, ..., N,\ t = 1, ..., T$ where y_{it} denotes the response variable of interest, i.e. it lies in the interval $[0, 1]$. Importantly, it may take on the values of 0 or 1. The whole analysis is based with the logistic specification of the conditional expectation function (CEF) for the first moment of y_{it}:

$$\mathbb{E}[y_{it}|\boldsymbol{x}'_{it}\boldsymbol{\beta}, \alpha_i] = \frac{\exp(\alpha_i + \mathbf{x}'_{it}\beta)}{1 + \exp(\alpha_i + \mathbf{x}'_{it}\beta)} \equiv \Lambda_{it} \tag{3.2}$$

After some transformations working with a pairwise of outcome variables, I arrive at my conditional moment:

$$\frac{1}{N}\sum_{i=1}^{N}[y_{i1}(1 - y_{i2}) - \exp((\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\beta)y_{i2}(1 - y_{i1})|\mathbf{x}'_{i1}, \mathbf{x}'_{i2}] = 0 \tag{3.3}$$

where a range of different instrumental variables could be used to identify β. Base on the moment condition where β is multiplied with $(\mathbf{x}'_{i1} - \mathbf{x}'_{i2})$, I will use this term as the instrument. My estimator would hence be the $\hat{\beta}$ that solves the unconditional moment of:

$$\frac{1}{N}\sum_{i=1}^{N_{12}}[y_{i1}(1 - y_{i2}) - \exp((\mathbf{x}'_{i1} - \mathbf{x}'_{i2})\beta)y_{i2}(1 - y_{i1})]\,(\mathbf{x}'_{i1} - \mathbf{x}'_{i2}) = 0 \tag{3.4}$$

The key extension is to adopt the multinomial-logit based conditional expectations for $y_{it}^{[k]}$:

$$\mathbb{E}[y_{it}^{[k]}|\mathbf{x}'_{it}, \alpha_i] = \frac{\exp(\alpha_i^{[k]} + \mathbf{x}'_{it}\beta^{[k]})}{1 + \sum_{k=1}^{K}\exp(\alpha_i^{[k]} + \mathbf{x}'_{it}\beta^{[k]})} \equiv \Lambda_{it}^{[k]} \quad \text{where } k = 1, \dots, K$$

Through similar transformation as in the univariate case, I have the analogue of the moment condition for the multinomial setting:

$$\frac{1}{N}\sum_{i=1}^{N}\left[y_{it}^{[k]}(1-\sum_{j=1}^{K}y_{is}^{[j]})-\exp[(\mathbf{x'}_{it}-\mathbf{x'}_{is})\beta^{[k]}]y_{is}^{[k]}(1-\sum_{j=1}^{K}y_{it}^{[j]})\right]=0$$

where it requires a pairwise non-autocorrelation between subgroups in period t and period s, i.e. $\mathbb{E}[\epsilon_{is}^{[k]}\epsilon_{it}^{[j]}|\mathbf{x}'_{it},\alpha_i]=0$.

As a choice probability model, the multinomial logit bears the inherent feature that discrimination among alternatives reduces to a series of pairwise comparisons which are unaffected by the characteristics of alternatives other than the pair under consideration. This well-known limitation of the standard multinomial logit, usually termed independence of irrelevant alternatives (IIA), implies zero correlation between the disturbances of the utilities associated with the various alternatives. In the fractional context, however, an analogous statement of independence between the errors of shares equations is not consistent with the unit-sum nature of the responses. By construction, the use of multinomial logit as the conditional mean modelling tool shows that the ratio between the conditional means of two different components of y_{it} functionally independent from characteristics of alternatives other than the considered pair.

The motivation for the mixed logit model arises from these limitations of the standard logit model. In this next subsection, I will discuss the mixed logit in its original form before combining with my estimators.

3.2.2 Mixed logit

I begin by briefly discussing the generic structure of discrete choice models. Different distributional specifications for β_i and ϵ_{it} generate different empirical models. One of the most widely used models is the conditional logit which arises when $\beta_i = \beta$ for $\forall n$ and each ϵ_{it} is an independent and identically distributed (iid) draw from the type I extreme value distribution.

Consider a simple multinomial logit y_{it} that could take K alternatives $k = 1, \ldots, K$. The probability that individual i chooses alternative k at time t takes the well known logit form (McFadden, 1974) of:

$$\Pr[y_{it} = k | x_{it}, \alpha_i] = \frac{\exp(x_{it}\beta^{[k]} + \alpha_i)}{\sum_{m=1}^{K} \exp(x_{it}\beta^{[m]} + \alpha_i)} \tag{3.5}$$

The conditional logit model embodies the IIA property which means that the odds ratio for any two alternatives is unaffected by the inclusion of any third alternative. To see this, consider the ratio of probabilities for alternatives k and l:

$$\frac{\Pr[y_{it} = k | x_{it}, \alpha_i]}{\Pr[y_{it} = l | x_{it}, \alpha_i]} = \frac{\frac{\exp(x_{it}\beta^{[k]} + \alpha_i)}{\sum_{m=1}^{K} \exp(x_{it}\beta^{[m]} + \alpha_i)}}{\frac{\exp(x_{it}\beta^{[l]} + \alpha_i)}{\sum_{m=1}^{K} \exp(x_{it}\beta^{[m]} + \alpha_i)}} = \frac{\exp(x_{it}\beta^{[k]})}{\exp(x_{it}\beta^{[l]})} \tag{3.6}$$

As $\Pr[y_{it} = k | x_{it}, \alpha_i]$ and $\Pr[y_{it} = l | x_{it}, \alpha_i]$ share the same denominator, the observable attributes for all other alternatives drops out, and thus the odds ratio will not change with the addition of any third alternative. This property of the conditional logit model is a direct result of the independent type I extreme value assumption.

For fractional response, however, I do not make a discrete choice out of a set of alternatives, I instead have multiple proportions or shares with each of them being a fraction and adding up to unity. Therefore, in the fractional framework, the mixed conditional mean of each category reflects the weighted average of the logit formula evaluated at different values of the random coefficients, with the weights given by the predetermined density. The expression of these mixed conditional means do not depend on the values of β_i but rather the set of underlying parameters θ which underpin the distribution of β_i and also become my parameters of interest.

The probability densities for β_i can be specified with either a continuous or discrete mixing distribution. With a continuous mixing distribution $f(\beta_i|\theta)$, integrating over the domain of $\beta_i^{[k]}$ gives:

$$\Pr[y_{it} = k|x_{it}, \alpha_i] = \int_{\beta_i} \Pr[y_{it} = k|x_{it}, \alpha_i, \beta_i] f(\beta_i^{[k]}|\theta) d\beta_i^{[k]} \tag{3.7}$$

When the dimension of $\beta_i^{[k]}$ is moderate to large, analytical or numerical solutions for the above integral are generally not possible. However, $\Pr[y_{it} = k|x_{it}, \alpha_i]$ can be approximated via simulation (McFadden and Ruud 1994). This involves generating several random draws from $f(\beta_i^{[k]}|\theta)$, calculating $\Pr[y_{it} = k|x_{it}, \alpha_i, \beta_i^{[k]}]$ for each draw, and then averaging across draws. By the law of large numbers, this simulated estimate of $\Pr[y_{it} = k|x_{it}, \alpha_i]$ will converge to its true value as the number of simulations grows sufficiently large.

3.3 Fractional mixed logit

In my fractional framework, the same interpretation follows for my mixed conditional mean. The IIA for fractional model is interpreted as the assumption where the ratio of two conditional means between two categories are independent from any attributes of other categories. This is nonetheless not plausible in many cases. For example, in time allocation literature, the ratio of one's work time share relative to leisure time share will not be independent from whether he or she has to spend time for childcare. The presence or absence of child care time share would be expected to change the relative share between the other two groups.

Mixed logit signifies the importance of conditioning for correlated individual heterogeneity. This feature coincides with the key advantage of a panel data. In many studies, this longitudinal characteristic makes it viable in capturing unobservable individual heterogeneity. It has also played at a central theme in designing my estimator. The progression to mixed logit, hence, allows my framework to be complete with individual heterogeneity is now being dealt with at both intercepts (α_i) and slopes level (β_i) Consequently, I fully characterise the impacts of unobserved individual effects in this non-linear panel framework.

These advantages, however, come with a larger complication of multiple integration of logistic functions. Similarly to probit, although mixed logit has been known for many years, it has only become fully applicable since the recent advancement of simulation, both in terms of computing power and randomisation technique in simulation. Due to the computationally intensive integration that is inherent in mixed logit, early applications on customer-level data, such as Train et al. (1987) and Ben-Akiva, Bolduc, and Bradley (1993), included only one or two dimensions of integration, which could be calculated by quadrature. Improvements in computer speed and in the understanding of simulation methods have allowed the full power of mixed logit to be utilized. Among the studies to evidence this power are those by Bhat (2000) and Brownstone and Train (1999) on cross-sectional data, and Erdem (1996), Revelt and Train (1998), and Bhat (2001) on panel data.

In my implementation for fixed-effects fractional estimator, I encounter further difficulties. In addition to the multidimensional integral over the mixing distribution, I also have a quasi-differencing expression due to the fixed-effects elimination transformation. As the integral does not have a closed-form, it must be evaluated numerically. In practice, the integral has often been approximated though the simulated method of moments (SMM) using a number of R random draws from the mixing distribution. While a standard R ranging between 100 to 1000, a larger number of draws is usually needed to assure reasonably low simulation error in the estimated parameters. However, a larger number of draws translates into longer computer run-times (Train (2000)). Numerous procedures have been proposed in the numerical analysis literature for taking more efficient draws from a distribution rather than random ones (Sloan and Wozniakowski (1998), Morokoff and Caflisch (1995)).

These procedures attempt to reduce run times, also the simulation error that is associated with a given number of draws. One of them has been particularly useful for mixed logit estimation - the Halton sequence. This technique largely bases on dividing the unit interval into equal distances using prime numbers. As a result, when I invert these proportions into a density, it covers the shape of the distribution more evenly and efficiently compared to the standard method of random draw. Bhat (2001) showed that by using this technique with $R = 100$, the author achieved a lower simulation error in the estimated parameters than that of 1000 random numbers. I hence implemented this Halton draw procedure in my application.

In practice, the limitation with the continuous mixed logit model is the pre-specified mixing distribution which often takes an arbitrary parametric form. Although the mean estimates were similar across different distribution, the standard deviations varied substantially depending on the mixing specification (Revelt and Train (1998)). Hence, a non-parametric approach is provided by setting a discrete or step function distributions that can readily account for different features of the data. For the fractional framework, I employ both parametric and non-parametric methods that are used in the discrete literature. Our aim is to flexibly measure the distribution or step density of β_i. I focus on estimating the spread of the random parameter.

Our conditional mean is now

$$\mathbb{E}[y_{it}^{[k]}|x_{it}, \alpha_i^{[k]}, \beta_i^{[k]}] = \frac{\exp(x_{it}\beta^{[k]} + \alpha_i^{[k]})}{1 + \sum_{m=1}^{K-1} \exp(x_{it}\beta^{[m]} + \alpha_i^{[m]})} \tag{3.8}$$

where I model the individual heterogeneity at the parameter level. Therefore, the characteristics of β_i's distribution or the density itself become the central objects of interest in estimation. This is particularly relevant in the context of individual choices as I allow heterogeneous preferences among decision makers.

The mixed choice probabilities curtails the restrictive forecasting patterns of standard logit in discrete model as it does not exhibit the independence from irrelevant alternatives (IIA). In choice theory, the IIA assumption says that when people are asked to choose among a set of alternatives, their odds of choosing A over B should not depend on whether some other alternative C is present or absent. As this condition does not usually hold, it becomes one of the shortcomings for the standard multinomial logit. Mixed logit then relaxes this assumption and provides the unrestricted substitution pattern between choice alternatives. The ratio of mixed logit probabilities between two choices now depends on all the attributes of alternatives other than the involving pair itself. Mathematically, the denominators of the logit formula are now inside the integrals and therefore do not cancel with each other. The relative probability of choosing two different choices are no longer independent from other alternatives.

3.4 Econometric approach

In this section, I present two different approaches to estimate the distribution of the random coefficient β_i, both parametrically and non-parametrically. Specifically, I aim to measure the standard deviation for the distribution of β_i. In the parametric case, I use the simulated methods of moments to estimate mean and standard deviation of a predetermined distribution of β_i parametrically. For the non-parametric case, I use an evenly spaced grid to estimate the step density of β_i at each support point.

First, I have a multinomial fractional panel dataset which includes K fractional responses of $y_{it}^{[k]}$ and a covariates of $\mathbf{x}_{it}'$ where $y_{it}^{[k]}$ is the k^{th} category $(k = 1, ..., K)$ of the fractional responses for observation $i = 1, ..., N$ at time $t = 1, ..., T$. By construction, each $y_{it}^{[k]}$ lies in the unit interval of $[0, 1]$, adds up to 1 ($\sum_{k=1}^{K} y_{it}^{[k]} = 1$) and more importantly, they could take values at the boundary with non-trivial masses.

I start from the functional form of the expected conditional mean for $y_{it}^{[k]}$:

$$\mathbb{E}[y_{it}^{[k]}|\mathbf{x}_{it}', \alpha_i^{[k]}, \beta_i^{[k]}] = \frac{\exp(\alpha_i^{[k]} + \mathbf{x}_{it}'\beta_i^{[k]})}{1 + \sum_{m=1}^{K-1} \exp(\alpha_i^{[m]} + \mathbf{x}_{it}'\beta_i^{[m]})} \equiv \Lambda_{it}^{[k]} \quad \text{where } k = 1, \ldots, K-1 \tag{3.9}$$

with $\mathbb{E}[y_{it}^{[K]}|\mathbf{x}_{it}', \alpha_i^{[K]}, \beta_i^{[K]}] = \frac{1}{1+\sum_{m=1}^{K-1} \exp(\alpha_i^{[m]}+\mathbf{x}_{it}'\beta_i^{[m]})}$ be the baseline category.

Let $\epsilon_{it}^{[k]}$ be the conditional expected function (CEF) error for $y_{it}^{[k]}$, I can express:

$$y_{it}^{[k]} = \Lambda_{it}^{[k]} + \epsilon_{it}^{[k]} \quad \text{where } k = 1, \ldots, K-1 \tag{3.10}$$

It is noted that, until this point, the only thing that differentiates this from the standard multinomial logit model is that I am working with continuous fractional variables which have logistic conditional means. Therefore, I do not have choice probabilities. In addition, I only "mix" when I have cleaned the incidental parameters $\alpha_i^{[k]}$, rather than applying the integration onto the probabilities straight-away.

Mathematically, in the discrete choice framework, the conditional probability would be integrated out as:

$$\begin{aligned} \Pr[y_{it} = k|\mathbf{x}_{it}', \alpha_i^{[k]}, \beta_i^{[k]}] &= \frac{\exp(\alpha_i^{[k]} + \mathbf{x}_{it}'\beta_i^{[k]})}{1 + \sum_{k=1}^{K-1} \exp(\alpha_i^{[k]} + \mathbf{x}_{it}'\beta_i^{[k]})} \equiv \Lambda_{it}^{[k]} \\ \Rightarrow \Pr[y_{it} = k|\mathbf{x}_{it}', \alpha_i^{[k]}, \theta^{[k]}] &= \int_{\beta_i^{[k]}} \left(\Lambda_{it}^{[k]}\right) f(\beta_i^{[k]}|\theta^{[k]}) d\beta_i^{[k]} \end{aligned} \tag{3.11}$$

where $k = 1, \ldots, K-1$ and $f(\beta_i^{[k]}|\theta^{[k]})$ is the density of the mixing distribution for $\beta_i^{[k]}$.

For my fractional response, I instead keep the CEF of $y_{it}^{[k]}$ conditioning on $\beta_i^{[k]}$:

$$\mathbb{E}[y_{it}^{[k]}|\mathbf{x}'_{it}, \alpha_i^{[k]}, \beta_i^{[k]}] = \frac{\exp(\alpha_i^{[k]} + \mathbf{x}'_{it}\beta_i^{[k]})}{1 + \sum_{k=1}^{K-1} \exp(\alpha_i^{[k]} + \mathbf{x}'_{it}\beta_i^{[k]})} \equiv \Lambda_{it}^{[k]} \tag{3.12}$$

where $k = 1, \ldots, K-1$. As mentioned before, the reason why I do not mix is due to the presence of the individual fixed effects $\alpha_i^{[k]}$. I wish to clean them first before integrating the moment condition. I carry the same steps in the last chapter in order to reach a similar moment condition:

$$\mathbb{E}\left[y_{it}^{[k]}(1 - \sum_{l=1}^{K} y_{is}^{[l]}) - y_{is}^{[k]}(1 - \sum_{l=1}^{K} y_{it}^{[l]}) \exp(\Delta \mathbf{x'}_{its}\beta_i^{[k]}) \middle| \mathbf{x}'_{it}, \alpha_i^{[k]}, \beta_i^{[k]} \right] = 0 \tag{3.13}$$

At this stage, I take two different approaches to describe the distributional effect in β_i. The first is to presume a parametric specification for the mixing distribution and hence, estimate the corresponding parameters. In practice, the variation of $\beta_i^{[k]}$ could be studied by assuming a mixing distribution which could be continuous (normal, log-normal, uniform) or discrete and non-parametric. The main incentive is to allow individual differences in $\beta_i^{[k]}$ and hence capturing the heterogeneity in preferences and substitution patterns. I do not aim to restrict the form of the underlying distribution. Therefore, in the second approach, I apply an evenly spaced grid in order to measure the density mass of the random coefficients at these grid points. Although they are methodologically different, they both represent the same goal of capturing the size of the individual heterogeneity in the model coefficients.

3.4.1 A parametric approach

In this subsection, I predetermined a distribution of $f(\beta_i^{[k]})|\theta^{[k]}$ for $\beta_i^{[k]}$. Hence, instead of estimating $\beta_i^{[k]}$, my focus is to measure the underlying parameters $\theta^{[k]}$. As I do not observe $\beta_i^{[k]}$, I need to integrate equation (3.13) over the mixing distribution of $\beta_i^{[k]} \sim f(\beta_i^{[k]}|\theta^{[k]})$. It is noted in this step that I do not have any incentive to eliminate the $\beta_i^{[k]}$ as I have done with the α_i because they are the parameters of interest, not the incidental parameters.

I use two different techniques of integral approximation: the Simulated Method of Moments (SMM) and the Gauss-Hermite Quadrature (GHQ). My aim is to observe if there is any significant computing advantages from one or the other. As this task requires heavy computational power, its efficiency is important to us. Subsequently, I also need to repeat the procedure with hundreds of Monte Carlo replication, hence, using a more efficient technique would help us trim down the running time significantly.

The method of simulated moments is a technique of numerical integration using random numbers. I approximate the integral in equation (3.25) by randomly drawing a set of R samples of $\beta_i^{(r)}, r = 1, \ldots, R$ from the mixing distribution of $N(\beta, \sigma_\beta^2)$. I then derive the average of the moments over these R draws:

$$\frac{1}{R}\sum_{r=1}^{R} \mathbb{E}\left[y_{it}(1-y_{is}) - y_{is}(1-y_{it}) \exp(\Delta x'_{its}\beta_i^{(r)}) \middle| x'_{it}, \alpha_i \right] = 0 \qquad (3.14)$$

Next, I approximate the population expectation $\mathbb{E}$ by using the analogous sample average over N individuals:

$$\frac{1}{R}\sum_{r=1}^{R}\frac{1}{N}\sum_{i=1}^{N}\Big[y_{it}(1-y_{is})-y_{is}(1-y_{it})\exp(\Delta x'_{its}\beta_i^{(r)})\Big]=0 \tag{3.15}$$

Equivalently, I have:

$$\frac{1}{N}\sum_{i=1}^{N}\left[y_{it}(1-y_{is})-y_{is}(1-y_{it})\times\frac{1}{R}\sum_{r=1}^{R}\exp(\Delta x'_{its}\beta_i^{(r)})\right]=0 \tag{3.16}$$

Continuing to the implementation phase, I rewrite my β_i as:

$$\beta_i=\beta+\sigma_\beta\epsilon_i \text{ where } \epsilon_i\sim N(0,1)$$

The representation makes constructing the program much more straight forward and intuitive. I could now set the two parameters of interest (β,σ_β) directly in the expression of the moments while the variation in β_i now being taken care of by the standard normally distributed ϵ_i. Subsequently, I import equation (3.16) into my codes for the Monte Carlo simulation.

In the second approximation technique, Gauss-Hermite quadrature provides another numerical evaluation for an integral of the expression

$$\int_{-\infty}^{\infty} f(y_{it}|x_{it},\beta_i)\phi(\beta_i|\beta,\sigma_\beta^2)d\beta_i \tag{3.17}$$

which cannot be solved by analytical methods. The starting point of this approach is the Jacobian transformation. I need to appropriately change the variable β_i to ν so that it brought the term into the following form

$$\int_{-\infty}^{\infty} h(\nu;y_{it},x_{it},\beta,\sigma_\beta^2)\exp(-\nu^2)d\nu \tag{3.18}$$

where ν represents the grid of G Gauss-Hermite evaluation nodes. In the following step, I use this G-point Gauss-Hermite quadrature to approximate the integral as a weighted sum of G nodes ν_g with its corresponding weights w_g:

$$\int_{-\infty}^{\infty} h(\nu; y_{it}, x_{it}, \beta, \sigma_\beta^2) \exp(-\nu^2) d\nu \approx \sum_{g=1}^{G} w_g h(\nu_g) \tag{3.19}$$

These nodes and weights are calculated using the Hermite polynomials $H_G(\nu_g)$ with $g = 1, \ldots, G$, where the node ν_g are the roots of the polynomial $H_G(\nu_g)$ and the associated weights are calculated as

$$w_g = \frac{2^{G-1} G! \sqrt{\pi}}{[G H_{G-1}(\nu_g)]^2}$$

As β_i follows $N(\beta, \sigma_\beta)$, I have

$$\begin{aligned} \phi(\beta_i | \beta, \sigma_\beta^2) &= \frac{1}{\sqrt{2\pi}\sigma_\beta} \exp\left[-\frac{1}{2} \left(\frac{\beta_i - \beta}{\sigma_\beta} \right)^2 \right] \\ &= \frac{1}{\sqrt{\pi}\sqrt{2}\sigma_\beta} \exp\left[- \left(\frac{\beta_i - \beta}{\sqrt{2}\sigma_\beta} \right)^2 \right] \end{aligned} \tag{3.20}$$

Let $\nu = \frac{\beta_i - \beta}{\sqrt{2}\sigma_\beta}$ or $\beta_i = \beta + \sqrt{2}\sigma_\beta \nu$. The Jacobian of the old and new variable ν is

$$d\beta_i = \sqrt{2}\sigma_\beta \quad d\nu$$

Therefore, I can transform the integral in (3.17) into (3.18) as follow

$$\begin{aligned}
&\int_{-\infty}^{\infty} f(y_{it}|x_{it},\beta_i)\phi(\beta_i|\beta,\sigma_\beta^2)d\beta_i \\
&= \int_{-\infty}^{\infty} f(y_{it}|x_{it},\beta_i)\frac{1}{\sqrt{\pi}\sqrt{2}\sigma_\beta}\exp\left[-\left(\frac{\beta_i-\beta}{\sqrt{2}\sigma_\beta}\right)^2\right]d\beta_i \\
&= \int_{-\infty}^{\infty} f(y_{it}|x_{it},\beta_i)\frac{1}{\sqrt{\pi}\sqrt{2}\sigma_\beta}\exp\left[-\nu^2\right]\sqrt{2}\sigma_\beta d\nu \\
&= \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{\pi}} f(y_{it}|x_{it},\beta_i)}_{h(\nu;y_{it},x_{it},\beta,\sigma_\beta)}\exp\left[-\nu^2\right]d\nu
\end{aligned}$$

Then the Gauss-Hermite approximation to the integral (3.18) is obtained as in (3.19). Substituting my moment condition into this expression I have:

$$\sum_{g=1}^{G} w_g \frac{1}{\sqrt{\pi}}\mathbb{E}\left[y_{it}(1-y_{is}) - y_{is}(1-y_{it})\exp(\Delta x'_{its}(\beta+\sqrt{2}\sigma_\beta\nu_g))\right] = 0 \quad (3.21)$$

Finally, the analogous sample average of the above expression is:

$$\begin{aligned}
&\sum_{g=1}^{G} w_g \frac{1}{\sqrt{\pi}}\frac{1}{N}\sum_{i=1}^{N}\left[y_{it}(1-y_{is}) - y_{is}(1-y_{it})\exp(\Delta x'_{its}(\beta+\sqrt{2}\sigma_\beta\nu_g))\right] = 0 \\
\Leftrightarrow\ &\frac{1}{N}\sum_{i=1}^{N}\sum_{g=1}^{G} w_g \frac{1}{\sqrt{\pi}}\left[y_{it}(1-y_{is}) - y_{is}(1-y_{it})\exp(\Delta x'_{its}(\beta+\sqrt{2}\sigma_\beta\nu_g))\right] = 0
\end{aligned} \quad (3.22)$$

I then use equation (3.22) for constructing my programming codes.

3.4.2 A non-parametric approach

In this approach, instead of assuming and then finding a parametric structure of the coefficient's distribution, I estimate the probability mass of random coefficients at evenly spaced support points. In another words, I estimate the step density function for β_i, leaving its underlying structure completely unspecified. I apply the technique proposed by Fox, Kim, Ryan, and Bajari (2011) used for discrete choice literature and modify it for my moment condition.

The key feature of this approach is that irrespective of the non-linear model used, the dependent variable would be linearly related with my parameters of interest after an approximation. Hence, instead of optimizing over the non-linear form, I compute the probability under each grid point as if it were the true parameter, and then find the linear mixture of those models that best approximates the actual data.

A simple example is used to elaborate this method. Instead of using discrete response, I adapt my fractional model onto the example directly. Suppose that an econometrician is interested in estimating a simple cross-sectional fractional logit model with a dependent variable y_i, a scalar random coefficient β_i and suppose that this model has a single independent variable, x_i.

$$\mathbb{E}[y_i|x_i, \beta_i] = \frac{\exp(x_i\beta_i)}{1 + \exp(x_i\beta_i)}$$

Furthermore, suppose that the random coefficient β_i is known to have support on the $[0, 1]$ interval, I fix a large but finite grid of S equally spaced points and suppose that the grid points take on the values of $\frac{1}{S}, \frac{2}{S}, \ldots, 1$. The parameters of my model are now θ_s which represents the probabilities that β_i takes value at each grid point, i.e. $\sum_{s=1}^{S} \theta_r = 1$ and $\theta_s \in [0, 1]$. It then follows that the empirical conditional mean of the dependent variable y_i can be approximated by the linear combination of θ_s:

$$\mathbb{E}[y_i|x_i, \theta_s] \approx \sum_{s=1}^{S} \theta_s \frac{\exp(\frac{s}{S}x_i)}{1+\exp(\frac{s}{S}x_i)} \equiv \sum_{s=1}^{S} \theta_s z_s \text{ where } \sum_{s=1}^{S} \theta_s = 1 \;\&\; \theta_s \in [0,1] \tag{3.23}$$

As I recall that the support points $\frac{s}{S}$ are fixed and known, given x_i, the term z_s is a known regressor. Hence, my estimator would be an inequality constrained least square estimate of equation (3.23):

$$\hat{\theta} = \arg\min_{\theta_s} \mathbb{E}\left[\left(y_i - \sum_{s=1}^{S} \theta_s z_s\right)^2\right] \text{ subject to } \sum_{s=1}^{S} \theta_r = 1 \;\&\; \theta_s \in [0,1] \tag{3.24}$$

According to Fox et al. (2011), the least-squares estimator has a unique solution as long as the vector of z_s has a full rank of S and the estimator is consistent under standard assumptions for least squares. Through my simulated results, it is shown that the full rank condition is crucial for identification and it is also dependent on the non-linear form of z_r. For some forms of non-linearity, the full rank condition is vulnerable and prone to be violated.

3.5 Monte Carlo simulation results

3.5.1 Monte Carlo set-up

I first start with a single fractional dependent variable $y_{it} \in [0,1]$ with $i = 1, \ldots, N$ and $t = 1, \ldots, T$ where N is the number of individuals and T is the number of time periods. The data generating processes (DGPs) for the covariates comprise of a single regressor x_{it} and a normally distributed fixed effects α_i that correlates with the regressor:

$$x_{it} \sim U(0,1), \quad i = 1, \dots, N, \; t = 1, \dots, T,$$
$$\alpha_i \sim N(x_{it}, 0.5).$$

For the dependent variable, I have two different DGPs. In the first DGP, I adopt the process from the first chapter where the dependent variable is generated from a binomial distribution with a logistic mean. This gives us a fractionalised discrete variable. The advantage is to guarantee that the dependent variable will (i) have values at both bounds and (ii) do not exceed the coherent unit interval. Thus, y_{it} is generated as the share of successes of a binomial random variable with success probability Λ_{it},

$$y_{it} = w_{it}/c, \text{ where } w_{it} \sim B(c, \Lambda_{it}).$$

The success probability is the mean of the fractional variable, hence y_{it}: $\mathbb{E}[y_{it}] = \Lambda_{it}$. I specify $\Lambda_{it} = \Lambda(\beta_0 + \alpha_i + \beta_i x_{it})$ where $\beta_i \sim N(\beta, \sigma^2_\beta)$. Notice that I do not have the superscript of $[k]$ in this simple case since I have $K = 2$, in which, there will be only 2 fractions: y_{it} and its baseline $1 - y_{it}$.

In addition, I also look at a second DGP, I loosely call it the conditional expectation DGP. As suggested by its name, I set:

$$y_{it} = \Lambda_{it} + \epsilon_{it}$$

where the conditional expectation function (CEF) errors ϵ_{it} is explicitly additively separated from the conditional mean. The reason for this second DGP is that the magnitude of variation in ϵ_{it} can then be changed to examine the sensitivity of the moment condition due to larger CEF errors. On the other hand, I could not do the same with the first DGP as I do not have a direct control on how ϵ_{it} would vary. The downside for this second process is that I no longer have (i) values at extreme bounds and (ii) the conditional mean has to be set at somewhere around midpoint so that it leaves sufficient room for the variation of ϵ_{it}.

My main interest lies in the estimation of β and σ_β, which I set the true values to be $\beta = 1$ and $\sigma_\beta = 0.5$. In my baseline DGP, I set the number of individuals $N = 10,000$ and the number of time periods $T = 2$. Before applying any approximations, I restate the moment condition of a pairwise quasi-difference between time period t and s for the mixed logit estimators in this univariate case:

$$\int_{\beta_i} \mathbb{E}\left[y_{it}(1 - y_{is}) - y_{is}(1 - y_{it}) \exp(\Delta x'_{its}\beta_i) \middle|\, x'_{it}, \alpha_i\right] \phi(\beta_i|\beta, \sigma^2_\beta) d\beta_i = 0 \tag{3.25}$$

where $i = 1, \dots, N;\ t, s = 1, \dots, T$ and $\Delta x'_{its} \equiv x_{it} - x_{is}$.

3.5.2 Results for the parametric approach

In Table 3.1, I present the results for the two approximation techniques - the SMM and the GHQ using the first DGP. For the former, I follow equation (3.16) choosing $R = 100$ random draws while for the latter, I implement equation (3.22) adopting a $G = 8$ point quadrature. The results are similar across the two methods. Although the estimates for σ_β (0.48 for SMM and 0.49 for GHQ) are relatively close to the true value of 0.5, those of β are about 9% biased away from the true value of 1. I also note that the convergence

Table 3.1. SMM vs. GHQ using the Binomial DGP

	SMM $R = 100$	GHQ $G = 8$
$\beta = 1$	0.9165 (0.1024)	0.9196 (0.1044)
$\sigma_\beta = 0.5$	0.4846 (0.2221)	0.4919 (0.2243)
Convergence rate	0.31	0.30
No. of observations		10,000
Simulation time (minutes)	10.1	5.8

R is the number of random draws from the specified mixing distribution for the SMM
G is the number of nodes for the GHQ

Table 3.2. GHQ using the Binomial DGP

	$N = 20,000$ $T = 2$ $\beta = 1$ $\sigma_\beta = 0.5$	$N = 10,000$ $T = 10$ $\beta = 1$ $\sigma_\beta = 0.5$	$N = 10,000$ $T = 2$ $\beta = 1$ $\sigma_\beta = 1$
$\hat{\beta}$	0.9397 (0.0289)	0.9127 (0.0356)	0.7291 (0.1777)
$\hat{\sigma_\beta}$	0.2088 (0.0888)	0.1578 (0.0945)	0.7243 (0.2332)
Convergence rate	0.42	0.31	0.29
Simulation time (minutes)	10.4	6.7	4.9

rate is quite low at about 30 %.

As the GHQ converges quicker and produces typically similar results with the SMM, from this point forward, I would use GHQ for extended simulations. In this range of trials, I increase the sample size N, number of time periods T and also set different true values of the parameters (β, σ_β). In Table 3.2 I have 3 separate cases: i) I increase the sample size N from 10,000 to 20,000, ii) increase the number of time periods T from 2 to 10 and iii) change the true value of σ_β from 0.5 to 1.

Table 3.3. GHQ using the Binomial DGP

No. of Reps	50	100	200
$\beta = 1$			
Mean	0.9117	0.9035	0.9087
	(0.0281)	(0.0302)	(0.0372)
Median	0.9144	0.9037	0.9067
IQR	(0.8922, 0.9338)	(0.8839, 0.9268)	(0.8799, 0.9273)
$\sigma_\beta = 0.5$			
Mean	0.1783	0.2106	0.2343
	(0.1053)	(0.0999)	(0.1305)
Median	0.1701	.2212	0.2313
IQR	(0.1393, 0.2423)	(0.1299, 0.2962)	(0.1394, 0.3392)
Conv. rate	0.35	0.44	0.43
No. of obs.	5000	5000	5000
Sim. time (ms)	3.2	4.9	9.5

I observe a clear deterioration in the performance of $\hat{\sigma_\beta}$, it appears that the estimates drift away from the true value as N and T increase. On the other hand, the estimates for β do not change and still about 9% off the target. Subsequently, when I change the true value of σ_β to 1, the estimators of both parameters struggle to reach the true values. These results show that the performance of GHQ under the first DGP is not stable, particularly in estimating the spread. Additionally, there is a consistent bias in estimating β which does not improve as N or T increases.

In the next table, I increase the number of replication in order to record if there is any improvement as there is more rep according to the law of large number. The outcomes show that the same bias persists in the estimates of the mean β at about nine per cent whereas there is a very little improvement of σ_β. The estimate of the standard deviation is still far away from the true value.

From these results, it indicates two behaviours. First, both of the approximation techniques are not able to estimate σ_β in a stable manner, they seem not to be able to handle larger data set either. The relative good performance it achieved at $N = 10,000$ could not be replicated at a higher number of observations. Second, the $\hat{\beta}$ is consistently about 9% biased and this does not disappear as I change my settings. At this stage, I suspect that the current DGP has relatively large CEF errors such that contaminate the moment condition evaluation. In order to directly manipulate the amount of CEF errors, I employ the second DGP which separates the logit conditional mean and the CEF error.

In this DGP of y_{it}, I first set the CEF errors standard deviation to be as minimal as 0.001, $\epsilon_{it} \sim N(0, 0.001^2)$ and then increase this σ_ϵ in the subsequent simulations. I would like to see how the moment condition would perform in a reasonably easy environment before larger and larger CEF errors are in presence.

As my original moment condition does not perform well in estimating the σ_β. I make a step backward and also look at simpler and easier moment conditions. I am back from my starting point with the infeasible moment where I assume that I could observe α_i and impose the conditional mean function directly. Although this scenario is impossible in practice, it is a good tool to check the validity of my moment in a simulated environment. As if this moment could not even produce unbiased estimates, it will signal that there are assumption violations in the process of generating data. Additionally, I also look at a modified moment condition, where I divide my moment by $y_{is}(1 - y_{it})$. There are two reasons behind including this form: first, with the current DGP used, I do not have to worry about bounds in the mean time, so by dividing $y_{is}(1-y_{it})$, I do not drop any observations. More importantly, this particular form somehow produced a better result for some of my initial trials. Hence, I are interested in observing its results in parallel with my original moment. Lastly, in addition to all of these, I also half the σ_β from 0.5 to 0.25 as I would like to see how things are in a most simple and easy case. As if it could not even perform in this setting, there is little purposes for the analysis to carry more difficult and complicated scenario.

Thus, this experiment consider the following moment conditions:

(i) the infeasible moment condition:

$$\mathbb{E}[y_{it} - \Lambda_{it}|\alpha_i, \beta, \sigma_\beta] = 0$$

(ii) the modified moment condition:

$$\mathbb{E}\left[\frac{y_{it}(1-y_{is})}{y_{is}(1-y_{it})} - \exp(\Delta x_{its}\beta_i)|\beta, \sigma_\beta\right] = 0$$

(iii) the actual moment condition:

$$\mathbb{E}[y_{it}(1-y_{is}) - y_{is}(1-y_{it})\exp(\Delta x_{its}\beta_i)|\beta, \sigma_\beta] = 0$$

The results are reported in Table 3.4 - , respectively.

Table 3.4, 3.5 and 3.6 show the simulation results across the three different moment conditions: the infeasible, the modified and the original moment respectively. For each table, I examine the performance of the estimators across the baseline and three other scenario where I alternatively increase N, T and the standard deviation of the CEF error σ_ϵ.

Table 3.4. Gauss-Hermite - Infeasible moment condition with $R = 100$ replications

	$N = 500$ $T = 2$ $\sigma_\epsilon = 0.001$	$N = 10,000$ $T = 2$ $\sigma_\epsilon = 0.001$	$N = 500$ $T = 20$ $\sigma_\epsilon = 0.001$	$N = 500$ $T = 2$ $\sigma_\epsilon = 0.05$
$\beta = 1$				
Mean	1.0038	1.0004	1.0003	1.0195
	(0.0116)	(0.0027)	(0.0041)	(0.0238)
Median	1.0010	1.0006	1.0001	1.0148
IQR	(0.9959, 1.0116)	(0.9988, 1.0022)	(0.9978, 1.0031)	(1.0024, 1.0325)
$\sigma_\beta = 0.25$				
Mean	0.3265	0.2565	0.2569	0.4461
	(0.1257)	(0.0485)	(0.0752)	(0.2072)
Median	0.3410	0.2610	0.2534	0.4483
IQR	(0.2385, 0.4013)	(0.2212, 0.2904)	(0.2072, 0.2989)	(0.2817, 0.6074)
Conv. rate	0.68	0.98	0.83	0.64
Sim. time (ms)	0.9	10.6	11.3	1.7

Table 3.5. Gauss-Hermite - Modified moment condition with $R = 100$ replications

	$N = 500$ $T = 2$ $\sigma_\epsilon = 0.001$	$N = 10,000$ $T = 2$ $\sigma_\epsilon = 0.001$	$N = 500$ $T = 20$ $\sigma_\epsilon = 0.001$	$N = 500$ $T = 2$ $\sigma_\epsilon = 0.05$
$\beta = 1$				
Mean	1.0175	1.0190	1.0197	1.0512
	(0.0267)	(0.0056)	(0.0068)	(0.0839)
Median	1.0162	1.0177	1.0196	1.0475
IQR	(0.9998, 1.0361)	(1.0154, 1.0219)	(1.0148, 1.0247)	(1.0015, 1.0961)
$\sigma_\beta = 0.25$				
Mean	0.4271	0.4298	0.4319	0.9206
	(0.1074)	(0.0236)	(0.0125)	(0.1510)
Median	0.4352	0.4289	0.4321	0.9454
IQR	(0.3626, 0.5038)	(0.4170, 0.4437)	(0.4238, 0.4419)	(0.8400, 1.0079)
Conv. rate	0.99	1.00	1.00	0.95
Sim. time (ms)	1.2	17.4	8.6	1.2

Table 3.6. Gauss-Hermite - Actual moment condition with $R = 100$ replications

	$N = 500$ $T = 2$ $\sigma_\epsilon = 0.001$	$N = 5,000$ $T = 2$ $\sigma_\epsilon = 0.001$	$N = 500$ $T = 20$ $\sigma_\epsilon = 0.001$	$N = 500$ $T = 2$ $\sigma_\epsilon = 0.05$
$\beta = 1$				
Mean	0.9978	0.9907	0.9909	0.9771
	(0.0273)	(0.0061)	(0.0062)	(0.0537)
Median	0.9971	0.9904	0.9895	0.9731
IQR	(0.9828, 1.0157)	(0.9865, 0.9935)	(0.9869, 0.9960)	(0.9412, 1.0166)
$\sigma_\beta = 0.25$				
Mean	0.2651	0.1097	0.0607	0.3361
	(0.1175)	(0.0304)	(0.0348)	(0.1584)
Median	0.2713	0.0990	0.0634	0.3013
IQR	(0.1819, 0.3322)	(0.0861, 0.1407)	(0.0462, 0.0785)	(0.2159, 0.4353)
Conv. rate	0.45	0.40	0.28	0.46
Sim. time (ms)	2.9	10.6	13.0	2.2

Overall, I observe consistent estimates for the mean β across all there cases. There is no longer a 9-10% downward bias persisted as in the binomial DGP. The estimates of β are all close to the true value of 1 and do not vary much when I change N, T or σ_ϵ. Hence, it appears that the level of noise from the CEF error plays a crucial role. Not only does it make it more difficult to estimate the standard deviation σ_β, but it also prevents us from reaching the true value for β. Only when I could directly control the amount of variation in the CEF that I could see unbiased estimates for my mean β.

On the other hand, the estimates of σ_β vary significantly across the three cases. In the infeasible setting where I assume to observe the values of the individual heterogeneity α_i, the performance was not so good at small sample size (small N at 500 and small T at 2). It improves significantly when I increase N and T to 5000 and 10 respectively. These results show that it requires much more data than I expected to be able to measure the spread of β_i precisely. The last column in Table 3.4 also restates this behaviour. As soon as I increase the σ_ϵ, even the infeasible moment struggles to pin down the true value of σ_β correctly. It rather needs even larger N and T to produce an unbiased estimate.

σ_β are biased when I use the modified and actual moments. It is also noted that the performance does not improve when I increase N or T. This implies that these moment conditions are not a sufficient restriction to underpin the value of σ_β. Similar to the infeasible setting, it even gets further from the true value for the case of higher σ_ϵ. The difference I could notice from the two setting is the stability of the modified moment. When I increase N or T, the values of estimates for σ_ϵ stay the same. Meanwhile, the actual moments drifted away significantly when I vary the sample size.

3.5.3 Results for the non-parametric approach

I employ the second DGP of the previous section to examine the performance of the estimator in this non-parametric setting. The main difference lies in the construction of β_i. Instead of being generated by a normal distribution, the random coefficients β_i would have the following four generating processes:

(i) discrete and symmetric

$$\beta_i = \begin{cases} 1 \text{ with } p = 0.25 \\ 2 \text{ with } p = 0.50 \\ 3 \text{ with } p = 0.25 \end{cases}$$

(ii) discrete and asymmetric

$$\beta_i = \begin{cases} 1 \text{ with } p = 0.10 \\ 2 \text{ with } p = 0.65 \\ 3 \text{ with } p = 0.25 \end{cases}$$

(iii) continuous and symmetric with $\beta_i \sim \mathcal{N}(2, 0.5^2)$; and

(iv) continuous and asymmetric with $\beta_i \sim \mathcal{X}^2(2)$

Through these four different DGP of β_i, I would like to see if the estimates of probability at each support point could reflect the true distribution of the DGPs. At this stage, all I are interested in is simply an estimated step function that could depict the underlying density. The support points are set at $s = 1, 2$ and 3. In the DGP of the dependent variable, I have $y_{it} = \Lambda_{it} + \epsilon_{it}$ and set the CEF errors to be $\epsilon_{it} \sim N(0, 0.001^2)$. I have one single regressor $x \sim \mathcal{U}(0, 1)$ with $N = 10,000$ and $T = 2$. It is noticed that I use a significantly larger N for the non-parametric approach.

I have the same three models that are analogous to the parametric moment conditions in the previous section:

(i) the infeasible moment condition:

$$\mathbb{E}\left[y_{it} - \sum_{s=1}^{3} \theta_s \frac{\exp(sx_{it})}{1 + \exp(\alpha_i + sx_{it})} |\alpha_i, \theta_s\right] = 0$$

(ii) the modified moment condition:

$$\mathbb{E}\left[\frac{y_{it}(1 - y_{iu})}{y_{iu}(1 - y_{it})} - \sum_{s=1}^{3} \theta_s \exp(s\Delta x_{itu}) |\theta_s\right] = 0$$

(iii) the actual moment condition:

$$\mathbb{E}\left[y_{it}(1-y_{iu}) - y_{iu}(1-y_{it})\sum_{s=1}^{3}\theta_s \exp(s\Delta x_{itu})|\theta_s\right] = 0$$

Tables 3.7 and 3.8 show the simulated results across all of the four scenario with Table 3.8 adds an a 4^{th} support point onto the grid. In Table 3.7, I also report the standard deviation of the estimates in parentheses.

The simulated results from Table 3.7 show that none of the configurations could work in all cases. While the infeasible and modified moment conditions perform better in the first three DGPs where β_i is generated from a discrete distribution and normal distribution, my actual moment achieves the closest shape to the true distribution for the last DGP. The standard deviation of the estimates from the simulation shows that for those that are close to the target, the variation is reasonably small. I further examine the results by increasing the number of grid points from 3 to 4 in Table 3.8. I recognise that with my non-linear functional form being an exponential, the full rank condition for the regressors is violated considerably quickly with the increase in S. When the number of support points increases to only $S = 4$, the results in Table 3.8 become destabilised. I observe numerous combination of (1,0,0,0) for the estimated density from the modified and actual moment condition. It shows that by adding another support point onto the search grid makes the modified and actual moment conditions struggle to pin down the probabilities. They all now give an estimated probability simplex of $\{1, 0, 0, 0\}$.

At this point, my concerns focus on the violation of the full rank condition. Not only do I have extremely highly correlated covariates, I could also

Table 3.7. Estimates of probability at support points for different DGPs of β_i

Support points	1	2	3
Discrete DGP of β_i with true value of	0.25	0.50	0.25
Infeasible	0.25	0.49	0.26
	(0.11)	(0.22)	(0.10)
Modified	0.22	0.55	0.23
	(0.10)	(0.25)	(0.12)
Actual	0.33	0.67	0.00
	(0.03)	(0.10)	(0.20)
Discrete DGP of β_i with true value of	0.10	0.65	0.25
Infeasible	0.10	0.65	0.25
	(0.05)	(0.21)	(0.10)
Modified	0.06	0.71	0.23
	(0.03)	(0.30)	(0.11)
Actual	0.00	1.00	0.00
	(0.00)	(0.00)	(0.00)
Continuous DGP of $\beta_i \sim \mathcal{N}(2, 0.5^2)$			
Infeasible	0.13	0.74	0.13
	(0.05)	(0.30)	(0.04)
Modified	0.12	0.76	0.12
	(0.06)	(0.32)	(0.06)
Actual	0.18	0.82	0.00
	(0.14)	(0.02)	(0.00)
Continuous DGP of $\beta_i \sim \mathcal{X}^2(2)$			
Infeasible	0.65	0.00	0.35
	(0.59)	(0.10)	(0.16)
Modified	0.00	0.00	1.00
	(0.00)	(0.00)	(0.00)
Actual	0.96	0.04	0.00
	(0.01)	(0.01)	(0.00)
Number of Replication			100

Table 3.8. Estimates of probability at support points for different DGPs of β_i

Support points	1	2	3	4
Discrete DGP of β_i with				
true value of	0.25	0.50	0.25	0
Infeasible	0.25	0.49	0.26	0.00
Modified	1.00	0.00	0.00	0.00
Actual	1.00	0.00	0.00	0.00
Discrete DGP of β_i with				
true value of	0.10	0.65	0.25	0
Infeasible	0.10	0.65	0.25	0.00
Modified	1.00	0.00	0.00	0.00
Actual	1.00	0.00	0.00	0.00
Continuous DGP of $\beta_i \sim \mathcal{N}(2, 0.5^2)$				
Infeasible	0.12	0.76	0.12	0.00
Modified	1.00	0.00	0.00	0.00
Actual	1.00	0.00	0.00	0.00
Continuous DGP of $\beta_i \sim \mathcal{X}^2(2)$				
Infeasible	0.73	0.00	0.00	0.27
Modified	1.00	0.00	0.00	0.00
Actual	1.00	0.00	0.00	0.00

parametrise them as a function of each other. Hence, it is likely that with my non-linear form of an exponential, this non-parametric method could not adopt my covariates at more than 3 support points and is prone to have difficulties in identifying the density estimates at each of these. I subsequently experiment with a number of different ways to reduce this multicollinearity such as smaller values for the covariates, however, the unstable performance remains. In sum, I found that not every non-linear function could work with the setting of Fox et al. (2011). Through simulation results, I showed that the full rank condition is vital for the identification of density at support points and that the functional form of z_s is important in determining if the method could perform in a consistent manner.

3.6 Conclusion

This chapter attempts to extend the fractional fixed effects estimator to random coefficients. I focus on the estimation of the underlying density that generate the random parameters, both parametrically and non-parametrically. I incorporate the mixed logit model from the discrete choice literature into my fractional fixed effects approach. I subsequently use different simulation techniques to approximate the integral associated with the mixing distribution. I found that the data generating process is very important and the simulated moment condition is very sensitive to the magnitude of disturbance, from σ_β and especially from the CEF error noise σ_ϵ.

I adapt a non-parametric approach also from the discrete literature by Fox et al. (2011) in order to estimate the step function of the density over support regions of evenly spaced grid points. In all cases, it appears that the task of measuring the 'spread', or the level of heterogeneity in the model parameter is non-trivial. Different to what seems to be a one-size-fits-all solution by Fox et al. (2011) for any non-linear mixed models, I found that the success in estimating the probability mass depends largely on the non-linear form of the covariates. It is due to the fact that the functional form could potentially make the full rank condition violated quickly with a increase in the number of grid points.

Overall, I extended the theoretical capability of the econometric tool developed in the previous study in handling the individual varying coefficients. As I emphasised the importance of individual heterogeneity in my context, it is a natural step to extend the assumed random structure of α_i onto the coefficients of interest β_i. It relaxes the condition that every individual faces the same effect of covariates and allows the interpretation of β_i to vary across persons. For the tractability of the model, in the parametric approach, I assume that the random coefficient would follow a normal distribution of mean β and standard deviation σ. I found that the simulated method of moments could identify the mean values, and hence pin down correctly the location of the β_i distribution. However, it struggled to achieve the same precision in estimating the spread parameter of σ_β.

In conclusion, the study achieved the following two key points: (i) confirming that the proposed econometric tool could identify the average value of the random coefficients in the case where they vary and follow a parametric distribution; and (ii) the limitation of methods of moments as they would not be able to accommodate extra parameters of interest easily.

Chapter 4

Household time shares cooperative structure - A cross correlation approach

4.1 Introduction

In this chapter, I want to explore a different set of assumptions, econometric settings and empirical approaches. First, I want to experiment a different position of where my error terms entering the conditional expectation function (CEF). Instead of having an error term outside of the logistic function, I will now include it inside. This will change the interpretations of these CEF errors. Second, I also experiment a different setting for my outcome variables where each member of the household will form a simplex, rather than for each household as a whole. Third, in terms of empirical approach, instead of estimating slope coefficients, I look at the cross correlations. I study couples' correlative structure in their time allocations across work and home activities. Different to studies that focus on estimating causal effects, this study aims to look at the cross-correlations for all outcome variables.

I focus on three different components driving the coordination in time allocations. I classify potential factors into the following categories: unobservable individual preferences, observable events, and unknown time varying factors. While it is fairly straight forward to measure the contribution of observable factors, it is less clear among the unknowns: individual preferences and time-varying residuals. I hence disaggregate them by adapting the Hausman et al. (1984) Poisson Fixed Effects (PFE) model. Subsequently, I derive the contribution towards time allocation correlative structure from these three different aspects.

The level of details in existing studies for time allocations depends on the particular disaggregation of non-market time used. One can be fairly broad-brush, as Biddle and Hamermesh (1990) were in dividing non-market time into sleep and non-sleep, and as were Stancanelli and Stratton (2014) in examining the price of domestic servants time along with each spouses time. One can be somewhat more detailed, as in Kooreman and Kapteyn (1987), choosing multiple uses of home time. The main point is that the literature stemming from those studies has produced numerous estimates of the extent of complementarity/substitutability among spouses activities. While the particular disaggregation that have been used have differed, it should be possible to meta-analyse this line of research in such a way as to infer the nature of household production/consumption in a family context. While the interdependencies in these time allocations represent an important issue, there is still little micro-level evidence on the actual magnitude of these joint structures. Therefore, I aim to fill this gap by examining extensively the correlations among partners' time uses across all time shares for paid employment, child care, housework and leisure.

Using the exponential functional form for the conditional mean, I adapt the count data based PFE model for this study. The key feature of the approach is the multiplicative separability for the fixed effects which allows the decomposition between unknown preferences and shocks. Unlike the previous two chapters where the fixed effects were cancelled out, I now estimate them and explore its correlative nature among partners' time uses. In addition, the decomposition also allows me to observe the coordination among the remaining unobservable time varying shocks. As these have been assumed away in the literature, concerns have been raised over its potential impact onto the main findings. As a result, while these components usually disappear in standard studies, they are estimated in this analysis and form the core of my correlative structure.

In my findings, individual preferences show strong correlations across many time shares. It signifies the similarity within each couple in terms of views towards gender's role and division of labour. I also document strong joint structure due to external events. Finally, the time varying unknowns do not appear to hold any correlative patterns. It subdues those concerns towards its role in determining correlations among the time share outcomes.

4.2 Theoretical framework

I use Ichino, Olsson, Petrongolo, and Thoursie (2019)'s Couple Optimisation Problem to illustrate the role of individual preferences in household time allocation. Although it does not have all the features reflected in our model such as panel setting or multiple grouping, it represents the key mechanism of how preferences could play a significant role in household time allocation's joint structure.

Households enjoy a home-produced good H and a market-produced good C and allocate spouses' time between market work and home production. The household good H is produced with a combination of spousal inputs according to the following constant elasticity of substitution specification:

$$H = \left[sH_m^{\frac{\beta-1}{\beta}} + (1-s)H_f^{\frac{\beta-1}{\beta}} \right]^{\frac{\beta}{\beta-1}} \tag{4.1}$$

where $j = \{m, f\}$ denotes spouses' gender and $H_j \in [0, 1]$ is the share of time devoted by spouse j to home production. s and $1-s$ are parameters related to the relative efficiency of spouses in home production and, most importantly, β represents the elasticity of substitution between spousal inputs. We interpret β as representing couples' preferences about the combination of spousal time in home production.

In the labour market, the productivity of one unit of time of spouse j is P_j . With perfectly competitive labour markets, wages for each spouse are equal to P_j and the associate earnings Y_j decrease with the share of time devoted to home production:

$$Y_j = P_j(1 - H_j) \tag{4.2}$$

Each spouse faces a tax schedule $T(Y_j)$, which may be progressive. Couples choose the optimal time allocation of spouses in order to maximize joint utility. Their optimisation problem is given by:

$$\begin{aligned} &\max_{H_m, H_f, C} U(H, C) \\ &\text{subject to } C \leq [Y_m - T(Y_m)] + [Y_f - T(Y_f)] \end{aligned} \tag{4.3}$$

where H and Y_j are defined as above.

The first order condition is:

$$\frac{s}{1-s}\left(\frac{H_m}{H_f}\right)^{-\frac{1}{\beta}} = \frac{(1-\tau_m)P_m}{(1-\tau_f)P_f} \tag{4.4}$$

The specified β hence represents couples' preferences on how spouses should contribute to the household production. Therefore, they determine how time allocation coordination is formed and the way partners co-locate their time shares.

In this study, I aim to illustrate the same aspect of how individual preferences may play a significant role in the cooperative process of time allocation. However, I provide an alternative approach to show this role of individual preferences. I estimate the correlations among time shares with individual preferences as one of three separate key parts that contribute to the joint structure. It reflects the extent to which couples attribute their own preferences onto their decision making, apart from external events and other unknown time-varying factors.

Let $\alpha_j^{[k]}$ represent the individual preference of member j in the household where $j \in \{m, f\}$ indicates whether the person is male or female and $k \in \{w, c, h\}$ indicates which time categories the preferences refer to. In the following matrix, I present the intended correlation structure that I would like to measure in Table 4.1. I measure all the pairwise correlations between individual preferences across both members of the household and three different time allocations. Due to construction which uses multinomial logistic for the conditional expectation, each member has to use one of their time use as the reference group and therefore leisure shares from both partners have been omitted. Hence, we could only observe three time allocations from each member.

Table 4.1. Correlations among individual preferences towards time uses

	Male			Female		
	Work	Child	House	Work	Child	House
Male						
Work	1					
Child	$c(\alpha_m^{[w]}, \alpha_m^{[c]})$	1				
House	$c(\alpha_m^{[w]}, \alpha_m^{[h]})$	$c(\alpha_m^{[c]}, \alpha_m^{[h]})$	1			
Female						
Work	$c(\alpha_m^{[w]}, \alpha_f^{[w]})$	$c(\alpha_m^{[c]}, \alpha_f^{[w]})$	$c(\alpha_m^{[h]}, \alpha_f^{[w]})$	1		
Child	$c(\alpha_m^{[w]}, \alpha_f^{[c]})$	$c(\alpha_m^{[c]}, \alpha_f^{[c]})$	$c(\alpha_m^{[h]}, \alpha_f^{[c]})$	$c(\alpha_f^{[w]}, \alpha_f^{[c]})$	1	
House	$c(\alpha_m^{[w]}, \alpha_f^{[h]})$	$c(\alpha_m^{[c]}, \alpha_f^{[h]})$	$c(\alpha_m^{[h]}, \alpha_f^{[h]})$	$c\alpha_f^{[w]}, \alpha_f^{[h]})$	$c(\alpha_f^{[c]}, \alpha_f^{[h]})$	1

4.3 Econometric approach

Instead of estimating the causal relationships between independent and dependent variables, my focus is the joint structure among the time shares. I study the contribution to the correlation among predicted outcomes from three different components. They represent three groups of factors that may drive the household joint time allocation: individual preferences, observables and unknown time-varying shocks. The approach is explained through a sequence of steps from constructing my outcome variables, adapting the PFE model, decomposing them into parts and finally, forming the correlation matrix.

First, the outcome variables are constructed from the same time use module within the longitudinal data of HILDA. The survey asked the respondents "How much time would you spend on each of the following activities in a typical week?". There were nine activities and we group them into four categories: paid employment, child care, housework and leisure which broadly includes the rest of available time.

Let $y^{[w]}_{it,j}, y^{[c]}_{it,j}, y^{[h]}_{it,j}$ and $y^{[l]}_{it,j}$ be the time shares an individual $j \in \{m, f\}$ of household $i = 1, \ldots, N$ at time $t = 1, \ldots, T$ spend on paid work (w), child care (c), home production (h) and leisure (l) respectively in a typical week. There are two constraints for these shares. First, they are all in the unit interval of $[0, 1]$. Second, they form a total of unity within each household member. The following conditions illustrate mathematically these constraints:

$$\begin{aligned} &0 \leq y^{[k]}_{it,j} \leq 1 \text{ for } k \in \{w, c, h, l\} \text{ and } j \in \{m, f\} \\ &y^{[w]}_{it,j} + y^{[c]}_{it,j} + y^{[h]}_{it,j} + y^{[l]}_{it,j} = 1 \text{ for } j \in \{m, f\} \end{aligned} \tag{4.5}$$

I utilise the same logistic function for these fractions, however, there are two main differences compared to how they were previously modelled in Chapter 2. First, the error term $\epsilon^{[k]}_{it,j}$ for the corresponding $y^{[k]}_{it,j}$ are now additive to the rest of the covariates $X_{it,j}$ and fixed effects $\alpha^{[k]}_j$, rather than staying outside of the logistic function as a CEF error. Here I keep the same notation for the individual j's fixed effects/preferences for time share k $\alpha^{[k]}_j$. Second, I group time shares within each member $j \in \{m, f\}$ rather than for the whole household i.

Choosing leisure time share as the baseline group for both male and female, the expression for time shares are as follow:

$$
\begin{aligned}
y_{it,j}^{[k]} &= \frac{\exp(\alpha_{i,j}^{[k]} + X_{it,j}'\beta_j^{[k]} + \epsilon_{it,j}^{[k]})}{1 + \sum_{n\in\{w,c,h\}} \exp(\alpha_{i,j}^{[n]} + X_{it,j}'\beta_j^{[n]} + \epsilon_{it,j}^{[n]})} \text{ for } k \in \{w,c,h\} \text{ and } j \in \{m,f\} \\
y_{it,j}^{[l]} &= \frac{1}{1 + \sum_{n\in\{w,c,h\}} \exp(\alpha_{i,j}^{[n]} + X_{it,j}'\beta_j^{[n]} + \epsilon_{it,j}^{[n]})} \text{ as the baseline for each } j \in \{m,f\}
\end{aligned}
\tag{4.6}
$$

Next, for each member, I divide each time shares of $y_{it,j}^{[w]}, y_{it,j}^{[c]}, y_{it,j}^{[h]}$ by $y_{it,j}^{[l]}$. This gives me an exponential function between the newly created outcome variables and the covariates:

$$
\frac{y_{it,j}^{[k]}}{y_{it,j}^{[l]}} = \exp\left[\alpha_{i,j}^{[k]} + X_{it,j}'\beta_j^{[k]} + \epsilon_{it,j}^{[k]}\right] \text{ for } k \in \{w,c,h\} \text{ and } j \in \{m,f\} \tag{4.7}
$$

The system of eight equations in (4.7) is similar in spirit to a seemingly unrelated regression (SUR) structure. The key feature that resemblances between my set of equation and a SUR is that the error terms $\epsilon_{it,j}^{[k]}$ are allowed to be correlated across the equations. This reflects the fact that household jointly determine the amount of time allocated to each activity across members.

Second, I adapt the PFE approach pioneered by Hausman et al. (1984) (HHG). PFE models were originally used for panel data when the outcome variable is count data. Hausman et al. (1984) proposed the method with an application for the number of patents filed by firms. In their empirical study, they wanted to control for the firm fixed effects and hence developed this approach. Wooldridge (1999) provided evidence that the PFE models maintain robustness properties as long as the conditional mean assumption of exponential form holds.

$$\mathbb{E}\left[\frac{y_{it,j}^{[k]}}{y_{it,j}^{[l]}}\middle|\alpha_{i,j}^{[k]}, X_{it,j}\right] = \exp\left[\alpha_{i,j}^{[k]} + X_{it,j}'\beta_j^{[k]}\right] \text{ for } k \in \{w, c, h\} \text{ and } j \in \{m, f\} \tag{4.8}$$

Third, I estimate the individual fixed effects. Similar to the linear function, while the estimated $\alpha_{i,j}^{[k]}$ are inconsistent, they are unbiased. Therefore, the same estimates for $\beta_j^{[k]}$ would be achieved either if dummy variables for individuals were included or HHG approach were used. In this chapter, I estimate these $\alpha_{i,j}^{[k]}$ using HHG approach.

At this stage, I have three different components of the predicted outcomes. These are multiplicatively separable and can present themselves with their own intuition. The first component is $\hat{\alpha}_{i,j}^{[k]}$ which represents the unobservable individual preferences. The second component is $X_{it,j}'\hat{\beta}_j^{[k]}$ which represents observable shocks and characteristics. The last component is $\hat{\epsilon}_{it,j}^{[k]}$ which serves as unknown time-varying shocks.

Fourth, I finally construct the estimated correlation matrix which indicates the attribution of different components onto the joint structure. There will be one table for each component. The interpretation depends on the sign, the statistical significance of the estimate, and whether it is a cross-correlation between two partners.

I focus on the cross-correlations between the two partners. There are three broad cases: i) If they are not statistical significant, indicated by the missing accompanied asterisk, they are deemed to be low to no correlation; ii) If the sign is positive, it indicates a co-movement in the same direction towards

Table 4.2. Correlations among partners' time allocations

	Male			Female		
	Work	Child	House	Work	Child	House
Male						
Work	1					
Child	$c\hat{G}_m^{[w]}, \hat{G}_m^{[c]})$	1				
House	$c(\hat{G}_m^{[w]}, \hat{G}_m^{[h]})$	$c(\hat{G}_m^{[c]}, \hat{G}_m^{[h]})$	1			
Female						
Work	$c(\hat{G}_m^{[w]}, \hat{G}_f^{[w]})$	$c(\hat{G}_m^{[c]}, \hat{G}_f^{[w]})$	$c(\hat{G}_m^{[h]}, \hat{G}_f^{[w]})$	1		
Child	$c(\hat{G}_m^{[w]}, \hat{G}_f^{[c]})$	$c(\hat{G}_m^{[c]}, \hat{G}_f^{[c]})$	$c(\hat{G}_m^{[h]}, \hat{G}_f^{[c]})$	$c(\hat{G}_f^{[w]}, \hat{G}_f^{[c]})$	1	
House	$c(\hat{G}_m^{[w]}, \hat{G}_f^{[h]})$	$c(\hat{G}_m^{[c]}, \hat{G}_f^{[h]})$	$c(\hat{G}_m^{[h]}, \hat{G}_f^{[h]})$	$c(\hat{G}_f^{[w]}, \hat{G}_f^{[h]})$	$c(\hat{G}_f^{[c]}, \hat{G}_f^{[h]})$	1

time share allocation, i.e. complementarity; or iii) If the sign is negative, it implies a division of labour, i.e. substitutability.

Let $\hat{G}_j^{[k]}$ represent an estimated component G from one of the three factors, for individual $j \in \{m, f\}$ for time shares $k \in \{w, c, h\}$, I have the correlations among factor G of partners' time allocations in Table 4.2.

My aim is to observe the following patterns: i) How individual preferences correlate across different time allocations and partners; ii) How observable events contribute to the joint structure; and iii) If there are any correlations left over in the unobservable time-varying shocks.

4.4 Results

In the process of achieving the exponential conditional mean for my outcome variables, I have taken a transformation, shown in Equation (4.7). In this step, each of the member's time shares for work, child care and house work has been divided against his or her own leisure shares. This has brought the ratio of $\frac{y_{it,j}^{[k]}}{y_{it,j}^{[l]}}$ to the exponential form and hence, making it adaptive to the

Table 4.3. Correlation matrix - original data

		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	-0.0420*	1.0000				
	House	-0.0710*	0.3001*	1.0000			
Female	Work	0.0388*	-0.0200*	0.0460*	1.0000		
	Child	0.0237*	0.2326*	0.0122	-0.2216*	1.0000	
	House	0.0249*	0.0253*	-0.0104*	-0.2812*	0.2536*	1.0000

Notes: * indicates that the correlation is significant at 5%

PFE framework. This odds ratio transformation is often used when handling multinomial logistic form. Although it is a relatively standard procedure, I no longer have all of the four time shares from each partner from this step onwards. The subsequent interpretations are thus implied for the ratios of time shares. Therefore, before getting to the main result, I examine the raw time shares data and explore the correlative patterns across all four shares in Table 4.3.

In this table, there are two separate groups of correlations. First, there are sets of within correlation for each person's time allocations. They can be found at the upper left and the bottom right corners of the matrix. These usually show an expected sign of negative correlations between the work hours with the at-home time shares such as child care or house work. As seen from the Table 4.3, the work hours strongly correlated negatively with the other two time shares. It applies for both males and females.

Second, the bottom left matrix shows cross-correlations between time shares of the female partner with their male counterparts. The interpretation for this sub-matrix is as follow. The main diagonal indicates the correlations between the partners' allocations with respect to the same type of time shares. Hence, if it shows a positive figure, it indicates a complementarity while a negative number would imply otherwise. Reversed with this, the off-diagonal correlations show the cross-group cross-partner relationships and hence it would indicate substitutability by a positive number. Intuitively, the main diagonal shows the tendency of partners doing the same thing together while the off-diagonal measures the likelihood of division of labour or specialisation.

For correlations within the same type of activity, I find that couples tend to allocate more work hours and especially, child care hours together (complementarity) whereas they would substitute on housework. The strongest correlation of 0.2326 is documented for their child care shares while the figures for work and house work are fairly small (0.0388 and -0.0104). It suggests that for childcare hours, there is a mutual benefit of both partners spending more time together.

For cross-category, the specialisation patterns appear for all pairs except for the negative correlation between males' child care and females' work hours. The evidence further illustrates the trade-off between work commitment and family responsibility that coupled family partners often need to face.

I now follow the steps stated in the previous section, transform the shares into odds ratios, apply PFE and subsequently decompose the predicted outcome into the three following components: the estimated fixed effects $\hat{\alpha}_{i,j}^{[k]}$ which represent personal preferences; observable covariates $X_{it,j}'\hat{\beta}_j^{[k]}$; and the unobservable time-varying residuals $\hat{\epsilon}_{it,j}^{[k]}$ which depicts unknown shocks. The transformation means that interpretations onto the joint structure is now with respect to the ratio of each time shares over its leisure share,

Table 4.4. Correlation - Estimated Individual Fixed Effects

		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	0.5503*	1.0000				
	House	0.5363*	0.6504*	1.0000			
Female	Work	0.0155	-0.0003	0.0652	1.0000		
	Child	0.1192*	0.3260*	0.1603*	0.1820*	1.0000	
	House	0.1053*	0.1562*	0.1000*	0.3068*	0.5834*	1.0000

Notes: * indicates that the correlation is significant at 5%

rather than the shares themselves. It, hence, may cause changes in within-correlations among time shares for each person. However, as I focus on the cross-correlation patterns between male and female partners, it does not change our key interpretations.

Individual preferences

The first component reflects the joint structure among partners' personal preferences on time use. As these fixed effects could be interpreted as personal preferences towards each type of time allocations, it reveals interesting pattern within households on how couples decide to share their work commitment and negotiate their household responsibilities. In Table 4.4, I find that the personal preferences towards childcare explain complementarity around the housework but not around paid work hours. It suggests that there is a matching pattern where partners tend to have similar preferences towards child care commitment and taking care of house work. However, they are flexible with each of their own choice on how many hours of work they should do. The figures also reveal that women's preference towards work hours has nothing to do with any of her man's view about his time allocations (0.02, -0.00 and 0.07). On the other hand, men's preference for work implies pos-

Figure 4.1. Individual preferences towards spending time for children

itive correlation, or substitutability, towards his female partners' preference for house work and child care hours.

Preferences towards spending time for child care are highlighted as the strongest correlation among all the cross-correlation between the partners. It shows that for couples who enter parenthood, there is a common desire to spend more time with their new member more than any other specialisations or substitutions. Figure 4.1 shows a significant positive correlation between males' and females' preferences towards spending time on childcare.

Through these results, it reaffirms that individual uniqueness sheds insights into the cooperative household time allocation.

Observables

The second components reflects changes and shocks that are observable and impact on the whole household time allocations. They include employ-

Table 4.5. Correlation matrix - Observables

		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	0.5775*	1.0000				
	House	0.3678*	0.6007*	1.0000			
Female	Work	-0.3384*	-0.0034	0.0071	1.0000		
	Child	0.2025*	0.3300*	0.1853*	-0.3152*	1.0000	
	House	0.2792*	0.1898*	0.0333	-0.4032*	0.0731*	1.0000

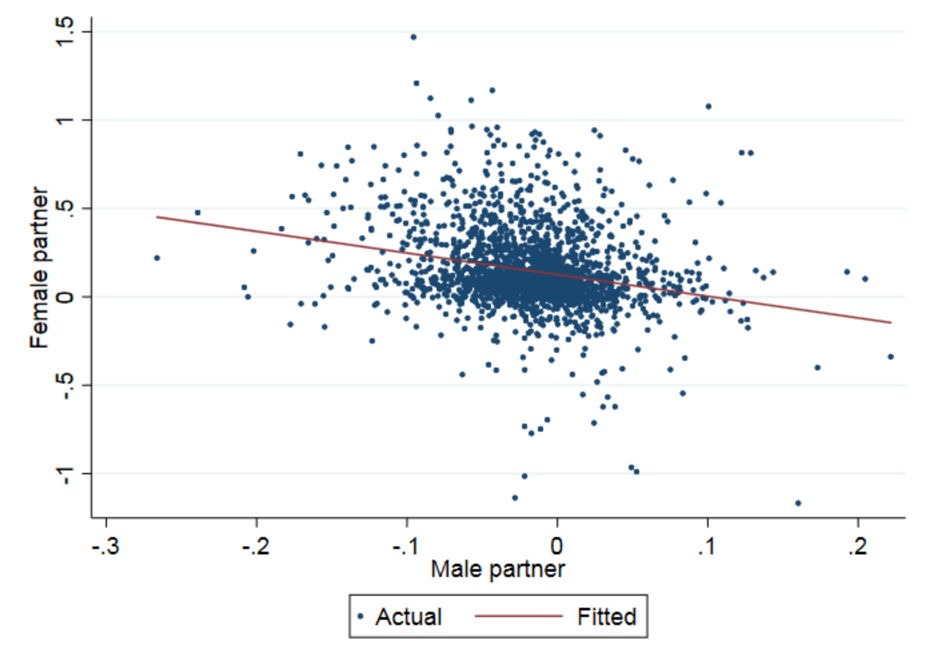
Notes: * indicates that the correlation is significant at 5%

ment status changes, state unemployment rate, family size and other sudden life events.

The figures from Table 4.5 illustrate the pattern of division of labour between male and female partners. The way the partners react to these observable shocks and changes aligns with our expectation. First, there is a significant negative correlation between males' and females' work hours. It implies a negotiating process between the two in which who would be able to keep his or her employment commitment whereas the other has to reduce or even withdraw from the labour market participation. Hence, rather than both make an attempt to reduce their market involvement and spend more time at home once facing these changes, it seems to be the case that they would decide to divide the works with one would handle the burden of housework while the other would be responsible for the financial condition of the household.

Besides, I also find an interesting difference between the way males and females shuffle their own time in the incident of facing a change. While men can increase their household commitment together with their paid work, certainly with an expense of reduced leisure and sleeping time, it does not

Figure 4.2. Individual time spent on paid work when facing observable changes



appear to be the case for women. The women appear to have to switch around their work commitment and houseworks with no room for doing both in the event of a shock or a change. This could potentially reflect that they often do not have a buffer from leisure in order to expand their capacity of active hours. Once there is a component that requires more of their time, other components within the active time group have to make a sacrifice, as shown on Figure 4.3.

In short, it shows interesting distinction of how couples react to changes and shocks happen to their families.

Unknown shocks

The last component is captured by the disturbance term that are unobservable and time variant. This is the unexplained portion of the outcome variable. I am interested in their behaviour due to the common concerns

Table 4.6. Correlation matrix - Residuals

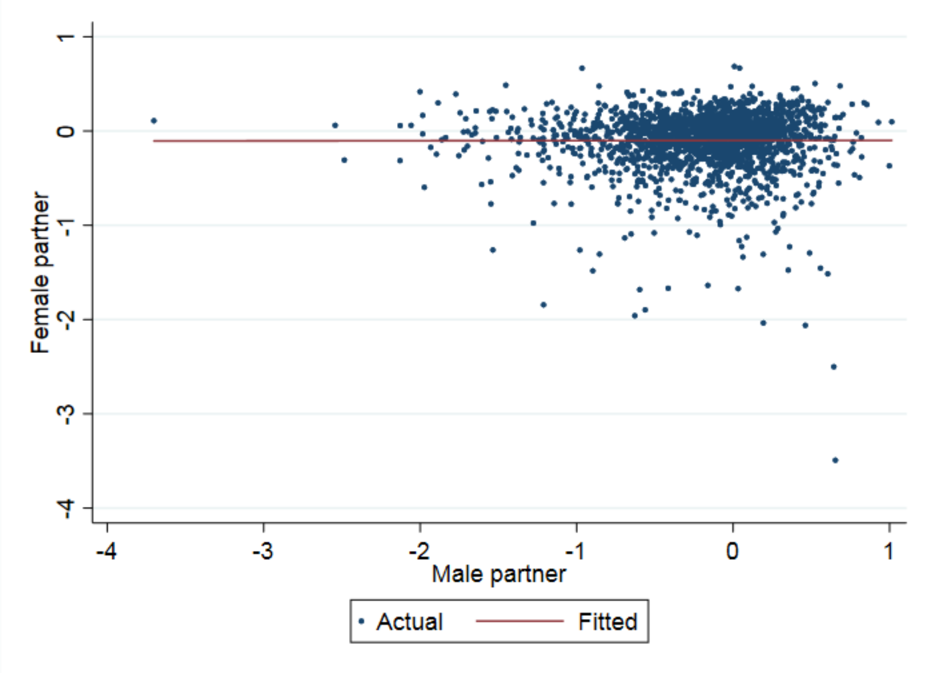
		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	0.2280*	1.0000				
	House	0.2133*	0.3264*	1.0000			
Female	Work	0.0141	0.0047	0.0206	1.0000		
	Child	-0.0041	0.0761	0.0158	-0.0089	1.0000	
	House	0.0217	0.0129	0.0090	0.0633	0.5634*	1.0000

Notes: * indicates that the correlation is significant at 5%

among the practitioners that it could produce biasedness in our estimation. Here in this application, I show that they do not explain any correlations between the partners' time shares and hence, controlling for the personal preference would capture most of the correlations in the joint structure among these shares. Consequently, it reaffirms the assurance towards potential critics that fixed effects method ignores the time-varying disturbance. The evidence is clearly shown in both Table 4.6 and Figure 4.3.

The main interest of the Table 4.6 is to examine if there is still any correlation left over inside this unobservable time-variant factors. As fixed effects have been long considered as an important factor need to be addressed in any panel data setting, concerns is still often raised with potential time-variant unobservables. The argument is that aside time-invariant heterogeneity, there could still be factors that changing over time and potentially correlate with our regressor of interest. In this study, I find that this concern is not an issue. I find no significant correlations across the time uses between the two partners. The fixed effects and observable covariates account for most, if not all, of the correlations across the time shares of the two partners. As we can clearly see form Figure 4.3, the unexplained shocks

Figure 4.3. Unexplained shocks in couples time allocation for housework



do not play any role in explaining the joint structure among household time allocations.

Finer Decomposition

Although the shocks component have been divided into observable and unobservable, the preferences have been grouped as a single component. In this part of the analysis, I further detach the preferences component into unobservable and observable. In my analysis, I would like to see if the results could show a difference of time allocation correlation across the two members coming from these panel averages versus the temporal deviation from these averages.

Table 4.7. Correlation matrix - Unobservable Preferences

		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	0.2988*	1.0000				
	House	0.1592*	0.4160*	1.0000			
Female	Work	0.0204*	0.0423*	0.0522*	1.0000		
	Child	0.0908*	0.2132*	0.0295*	0.1662*	1.0000	
	House	0.0891*	0.0957*	0.0678*	0.2631*	0.4796*	1.0000

Notes: *indicates that the correlation is significant at 5%

Table 4.8. Correlation matrix - Observable Preferences

		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	0.9238*	1.0000				
	House	0.8974*	0.8713*	1.0000			
Female	Work	-0.0300*	-0.0352*	0.0928*	1.0000		
	Child	0.2895*	0.5054*	0.0842*	-0.0454*	1.0000	
	House	0.2532*	0.3269*	0.1035*	-0.0154	0.6381*	1.0000

Notes: *indicates that the correlation is significant at 5%

In terms of preferences, there is an interesting difference in terms of work hours between the partners. While in Table 4.7, the unobservable parts suggest that partners tend to have complementarity in terms of work hours (significantly positive at 0.02), the observable preferences imply that they may want to substitute for work hours (significantly negative at -0.03) in Table 4.8.

Table 4.9. Correlation matrix - Observable Shocks

		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	0.4955*	1.0000				
	House	-0.0526*	0.0898*	1.0000			
Female	Work	-0.2479*	0.0462*	0.1635*	1.0000		
	Child	0.3014*	0.6285*	-0.2813*	-0.2715*	1.0000	
	House	0.2458*	0.1333*	0.2274*	-0.4130*	0.1493*	1.0000

Notes: *indicates that the correlation is significant at 5%

Table 4.10. Correlation matrix - Unobservable Shocks

		Male			Female		
		Work	Child	House	Work	Child	House
Male	Work	1.0000					
	Child	0.2444*	1.0000				
	House	0.2054*	0.4296*	1.0000			
Female	Work	-0.0104	-0.0101	0.0312	1.0000		
	Child	0.0010	0.0633*	-0.0035	0.1058*	1.0000	
	House	0.0030	0.0064	0.0176	0.1399*	0.5606*	1.0000

Notes: *indicates that the correlation is significant at 5%

In terms of time-varying shocks, there is expectedly a vast difference between observables and unobservables. Similar to the former results, as the observables contribute significantly to the joint structure of time allocations between the two partners while there is virtually no significant correlations documented in the unobservable shocks, as shown in Table 4.9 and Table 4.10.

4.5 Conclusion

In this chapter, I experimented with a different form and setting for the fractional responses and also the conditional expectation function error term. In terms of empirical approach, instead of pursuing causal estimates, I explored a cross-correlation method.

I focus on the correlative structure of household time allocation instead of any particular causal relationships. The analysis aims to detach multiple components of the outcomes that have different natural interpretations. In particular, I decompose the outcome into three different components: individual unique preferences $\alpha_{i,j}^{[k]}$; observable shocks and changes $X'_{it,j}\beta_j^{[k]}$; and finally the unexplained time variant errors $\epsilon_{it,j}^{[k]}$. I show that the last component does not explain any correlations between the two partners and hence, the standard practice of controlling for personal preferences should suffice. Consequently, opposite to standard critics towards fixed effects of leaving out time variant factors, the time variant components do not seem to hold significant correlations.

I found that individual preferences showed strong correlations across many time shares. It signified the similarity within each couple in terms of views towards gender's role and division of labour. I also documented strong joint structure due to external events. Finally, the time varying unknowns did not appear to hold any correlative patterns. It subdued concerns towards its role in determining correlations among the time share outcomes.

In terms of empirical evidence towards household time allocation, I find three interesting results. First, personal preference reveals patterns of selective matching among couples. They tend to match with respect to their preference towards taking care of children and housework. Second, there is a clear division of labour through observable changes and shocks that occur to the family rather than through partners' own preferences. Hence, it explains the negotiating process between the two members where they are optimising for the whole household rather than being able to do it alone. Third, I also find women have little or no room for increasing their active time spent for paid and unpaid works. Once there is a change or shock, they have to shuffle between the two. This is not the case for men where they can have ability to absorb the shock by reducing their leisure or rest time if needed.

Chapter 5

Conclusion

My thesis aims to study the importance of individual heterogeneity in a longitudinal non-linear setting. My approach highlights the importance of modelling natural boundedness of outcome variables through the application on time allocations. It is a comprehensive set of studies which took the advantage of the rich panel data from HILDA survey, developing a class of estimators that are capable of accommodating more aspects of the empirical study than existing tools. The "golden thread" of the econometrically important individual heterogeneity and empirically significant topic of time uses travel through my three key chapters.

In my first substantive chapter, I introduced the new class of estimators which can deal with more features of fractional panel data than existing methods. I then testified its performance through a range of Monte Carlo simulation before applying it to the empirical data of household time use. In terms of results, I found that women took a larger impact onto their time allocations when the couple entered parenthood. It showed that estimate of the negative impact onto their work hours from having a newborn baby were much larger compared to that of their male partner. Although the gap tended to reduce over time and also as the child grows up, it still persisted. I also found stronger impact of first born than any higher order births. Besides, couples with higher education, higher income, and especially, more progressive attitude towards gender's role appear to react stronger with the fathers allocate much more time shares than their peers in traditional families.

In the second study, I extended my estimation framework to allow for random coefficients. This exploration also allows me to add individual heterogeneity from intercept to the slope level. Although I found that none of the methods could precisely estimate the variation of the random coefficients β_i, I could document some interesting findings. First, it showed that the original method could identify the mean of β_i. Hence, it validates the performance of the proposed estimator in the case where coefficients were random. In addition, I found that the simulated moment condition was very sensitive to the magnitude of disturbance, from σ_β and especially from the CEF error noise σ_ϵ. My chosen functional form could also potentially made it even more difficult to trace down the shape of the mixing distribution.

In the third study, I experimented with a different form and setting for the fractional responses and also the conditional expectation function error term. Rather than aiming at finding causal estimates, I also explored a different approach using cross-correlations. I found that individual preferences showed strong correlations across many time shares. It signified the similarity within each couple in terms of views towards gender's role and division of labour. I also documented strong joint structure due to external events. Finally, the time varying unknowns did not appear to hold any correlative patterns. It subdued concerns towards its role in determining correlations among the time share outcomes.

Finally, as I complete this thesis during the COVID-19 pandemic lock-down, I find that the literature of time-use and household production function may give us a particularly interesting angle to study its impact. As many of us would be expected to stay at home or work from home in the months to come, it would be interesting to see how we will change our time allocations due to this shock and how much of our own individual preferences may play a role in the process of handling the shared work load. In addition, it also signifies the importance of understanding household production functions. While restaurants and other services may close due to the lock-down, consumption of goods and services may shift significantly from market to in-home.

Bibliography

Alesina, A., E. L. Glaeser, and B. Sacerdote (2005). Work and leisure in the u.s. and europe: Why so different? *NBER Macroeconomics Annual* 20, 1 – 100. 3

Argyrous, G. and S. Rahman (2017). How does paid work affect who does the childcare? an analysis of the time use of Australian couples. *Review of Economics of the Household* 15(2), 383 – 398. 4, 18

Ashenfelter, O. and J. Heckman (1974). The estimation of income and substitution effects in a model of family labor supply. *Econometrica* 42(1), 73 – 85. 2

Bauer, P. and L. Sonchak (2017). The effect of macroeconomic conditions on parental time with children: Evidence from the American time use survey. *Review of Economics of the Household* 15(3), 905 – 924. 4, 18

Baxter, J. and M. Bittman (1995). Measuring time spent on housework: A comparison of two approaches. *Australian Journal of Social Research* 1, 21 – 46. 11

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal* 75(299), 493 – 517. 2

Ben-Akiva, M., D. Bolduc, and M. Bradley (1993). Estimation of travel model choice models with randomly distributed values of time. *Transportation Research Record* 1413, 88 – 97. 71

Bhat, C. R. (2000). A multi-level cross-classified model for discrete response variables. *Transportation Research Part B: Methodological* 34(7), 567 – 582. 72

Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological* 35(7), 677 – 693. 72

Biddle, J. and D. Hamermesh (1990). Sleep and allocation of time. *Journal of Political Economy* 98(5), 922 – 943. 99

Blau, F. D. and L. M. Kaln (1996). Wage structure and gender earnings differentials: An international comparison. *Economica* 63(250), s29 – s62. 8, 16

Bonhomme, S. (2012). Functional differencing. *Econometrica* 80(4), 1337 – s1385. 10

Brownstone, D. and K. Train (1999). Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics* 89, 109 – 129. 20, 64, 72

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies* 47, 225 – 238. 10

Cherchye, J., B. De Rock, and F. Vermeulen (2012). Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. *American Economic Review* 102(7), 3377 – 3405. 3

Chiappori, P. and M. Mazzocco (2017). Static and intertemporal household decisions. *Journal of Economic Literature* 55(3), 985 – 1045. 18

Craig, L. and P. Siminski (2011). If men do more housework, do their wives have more babies? *Social Indicators Research* 101(2), 255 – 258. 4, 19

Department of Social Services and Melbourne Institute of Applied Economic and Social Research (2017). Household, Income and Labour Dynamics in Australia (HILDA) Survey - General Release 16 . 5

Elsas, R. and D. Florysiak (2015). Dynamic capital structure adjustment and the impact of fractional dependent variables. *Journal of Financial and Quantitative Analysis* 50(5), 1105 – 1133. 9, 20, 27

Erdem, T. (1996). A dynamic analysis of market structure based on panel data. *Marketing Science* 15, 359 – 378. 72

Foster, G. and L. S. Stratton (2018). Do significant labor market events change who does the chores? paid work, housework, and power in mixed-gender Australian households. *Journal of Population Economics* 31, 483 – 519. 4, 19

Fox, J. T., K. I. Kim, S. P. Ryan, and P. Bajari (2011). A simple estimator for the distribution of random coefficients. *Quantitative Economics* 2, 381 – 418. 81, 82, 95, 96

Gershuny, J. (2003). Changing times: Work and leisure in postindustrial society. *European Sociological Review* 19(2), 235 – 239. 3

Greene, W. H. (2002). Econometric analysis. *New York University* . 19, 39

Hamermesh, D. S. (2016). What's to know about time use? *Journal of Economic Surveys* 30(1), 198 – 203. 1, 4, 18

Hausman, J. A., B. H. Hall, and Z. Griliches (1984). Econometric models for count data with an application to the patents-r & d relationship. *Econometrica* 52, 909 – 938. 39, 99, 105

Hausman, J. A. and G. K. Leonard (1997). Superstars in the national basketball association: Economic value and policy. *Journal of Labor Economics* 15(4), 586 – 624. 9, 20, 24, 26

Heckman, J. J. (2015). Introduction to a theory of the allocation of time by gary becker. *The Economic Journal* 125(583), 403 – 409. 2

Hermalin, B. E. and N. E. Wallace (1994). The determinants of efficiency and solvency in savings and loans. *RAND Journal of Economics* 25(3), 361 – 381. 23

Herr, J. L. (2015). The labor supply effects of delayed first birth. *American Economic Review: Papers & Proceedings* 105(5), 630 – 637. 16

Ichino, A., M. Olsson, B. Petrongolo, and P. S. Thoursie (2019). Economic incentives, home production and gender identity norms. Technical report, IZA Discussion Paper. No. 12391. 100

Kan, M. Y. and S. Pudney (2008). Measurement error in stylized and diary data on time use. *Sociological Methodology* 38, 101 – 132. 11

Kling, C. L. and C. J. Thomson (1996). The implications of model specification for welfare estimation in nested logit models. *American Journal of Agricultural Economics* 78, 103 – 114. 64

Kooreman, P. and A. Kapteyn (1987). A disaggregated analysis of the allocation of time within the household. *Journal of Political Economy* 95(2), 223 – 249. 99

Kuznets, S. (1934). National income. Technical report, National Bureau of Economic Research. Research Bulletin No. 49. 2

Loudermilk, M. S. (2007). Estimation of fractional dependent variables in dynamic panel data models with an application to firm dividend policy. *Journal of Business & Economic Statistics* 25(4), 462 – 472. 9, 27

Lundborg, P., E. Plug, and A. W. Rasmussen (2017). Can women have children and a career? IV evidence from IVF treatments. *American Economic Review* 107(6), 1611 – 1637. 8, 16

Machado, P. M. (2004). A consistent estimator for the binomial distribution in the presence of incidental parameters: an application to patent data. *Journal of Econometrics* 119(1), pp. 73 – 98. 25

Maddala, G. S. (1991). A perspective on the use of limited-dependent and qualitative variables models in accounting research. *The Accounting Review* 66(4), 788 – 807. 22

McFadden, D. and K. Train (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15, 447 – 470. 64

Mincer, J. (1962). On-the-job training: Costs, returns, and some implications. *Journal of Political Economy* 70(5), 50 – 79. 2

Mitchell, W. (1912). The backward art of spending money. *American Economic Review* 2(2), 269 – 281. 2

Morokoff, W. J. and R. E. Caflisch (1995). Quasi-Monte Carlo integration. *Journal of Computational Physics* 122(2), 218 – 230. 72

Mullahy, J. (2015). Multivariate fractional regression estimation of econometric share models. *Journal of Econometric Methodology* 4(1), 71 – 100. 13

Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica* 16, 1 – 32. 9, 25, 31

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics* 11(6), 619 – 632. 8, 23, 24, 26, 29

Papke, L. E. and J. M. Wooldridge (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics* 145, 121 – 133. 9, 27

Podivinsky, J. M. and G. Stewart (2009). Modeling proportions: Random effects models of uk firm entry. *Singapore Economic Review* 54(3), 367 – 377. 9

Press, J. E. and E. Townsley (1998). Wives' and husbands' housework reporting: Gender, class, and social desirability. *Gender and Society* 12, 188 – 218. 11

Ramalho, E. A., J. J. S. Ramalho, and M. S. Coelho (2016). Exponential regression of fractional-response fixed-effects models with an application to firm capital structure. *Journal of Econometric Methods* 7(1), 1 – 18. 10, 27

Ramalho, E. A., J. J. S. Ramalho, and J. M. R. Murteira (2011). Alternative estimating and testing empirical strategies for fractional regression models. *Journal of Economic Surveys* 25, 19 – 68. 39

Reid, M. (1934). Economics of household production. Technical report, New York: Wiley. 2

Revelt, D. and K. Train (1998). Mixed logit with repeated choices: Households' choices of appliance efficiency level. *Review of Economics and Statistics* 80(4), 647 – 657. 64, 72, 73

Robinson and G. Godbey (1997). Economics of household production. Technical report, New York: Wiley. 3

Rogerson, R. and J. Wallenius (2019). Household time use among older couples: Evidence and implications for labor supply parameters. *The Quarterly Journal of Economics* 134(2), 1079 – 1120. 18, 45, 46

Sloan, I. and H. Wozniakowski (1998). When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *Journal of Complexity* 14, 1 – 33. 72

Stancanelli, E. and L. S. Stratton (2014). Maids, appliances and couples' housework: The demand for inputs to domestic production. *Economica* 81(323), 445 – 467. 99

Train, K. (2000). Halton sequences for mixed logit. Technical report, Department of Economics, University of California, Berkeley. Working Paper No. E00-278. 72

Train, K., D. McFadden, and M. Ben-Akiva (1987). The demand for local telephone service: A fully discrete model of residential calling patterns and service choice. *Rand Journal of Economics* 18, 109 – 123. 64, 71

Wagner, J. (2001). A note on the firm size export relationship. *Small Business Economics* 17(4), 229 – 237. 24, 25

Weischselbaumer, D. and R. Winter-Ebmer (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys* 19(3), 479 – 511. 8, 16

Winkelmann, R. and L. Xu (2019). Testing the binomial fixed effects logit model with an application to female labor supply. Technical report, Department of Economics, University of Zurich. Working Paper No. 321. 10, 28

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics* 90, 77 – 97. 105

Wooldridge, J. M. (2019). Correlated random effects models with unbalance panels. *Journal of Econometrics* 211(1), 137 – 150. 33