LABOR SUPPLY EFFECT OF CASH TRANSFER PROGRAMS: EMPIRICAL EVIDENCE FROM CHILD TAX CREDIT REFORM

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Yuanmeng Li

August 2023

ABSTRACT

Social protection programs are important element for the government's welfare policy and their socio-economic implications need to be carefully evaluated. In this paper, we considered the temporary reform of the Child Tax Credit(CTC) program in 2021 that transformed it into a nearly universal program. The quasi-experimental Intent-To-Treat estimator confirms that the change induces no negative responses. Apart from the baseline result, we use extended monthly data to show there exists a positive, significant, and delayed labor participation effect. By doing so we supply supports to the friction hypothesis to the delayed response. To explain the stark behavioral responses within a household, we estimated a structural econometric model of static games between the husband and the wife in a household. We found there to be a significant evidence for the presence of strategic interaction in this intra-household labor supply decisions.

BIOGRAPHICAL SKETCH

Yuanmeng Li graduated with a Bachelor of Science(Honors) in Economics and Mathematics of the first class from the University of St.Andrews in Scotland. He pursues a graduate degree in Applied Economics and Management at Cornell University.

From an early age, he has profound interests in a variety of subjects ranging from philosophy, politics, and history. He later focused on the Development Economics and issues such as inequality, social protection programs, and public policy. His interests in these issues inspired him to delve into further research in these subject matters and the development this thesis.

Upon completion of the Master of Science degree in Applied Economics and Management, he will join the Harvard Business School as a Research Associate in the Finance unit.

This document is dedicated to all Cornell graduate students.

ACKNOWLEDGEMENTS

This thesis would have not been possible without the support, help, guidance from a group of wise, kind, and caring individuals at Cornell and across the globe. It is only right to bid my sincerest gratitude to each and everyone of them.

I could never be grateful enough for the love and support from my family. Their consistent care throughout the twenty-four years of my life gave me the opportunity to write this thesis and have it presented. This paper is dedicated to you.

I would like to extend my gratitude to the Cornell intellectual community. First and foremost, I would like to thank the members of my special committee, Professor Ravi Kanbur and Professor George Jakubson for their invaluable input in this thesis. The quality and completion of this thesis benefited immensely from their insights, critics, and professional guidance. I would also like to thank other faculty members of Cornell University for their impact and help on this project and my development during my time at Cornell. Professor Douglas Miller and Professor Cynthia Lin-Lawell gave me the most generous support and constructive feedback in shaping and forming some parts of this thesis. My appreciation also goes to my cohorts and peers. I thank my old friends for being with me through ups and downs over the years and all that happened between us: excitement, encouragement, and even occasional quarrels. For the sake of auld lang syne. I thank my new friends for embarking on the this challenging journey with me. The bleak winters(about 5 months) of Ithaca have been a lot more colorful thanks all of you.

One's memory capacity is limited for all the kind individuals who have provided direct or indirect assistance and support intentionally or unintentionally

throughout this academic endeavor and my development along the way, but one's appreciation is with all of them. It is a real honor and pleasure to meet and connect with all those mentioned and not mentioned, their kindness and sagacious contributions will forever be remembered and cherished.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

President Biden delivered one of his campaign promises as the Child Tax Credit(CTC) enacted on 15th July 2021. The CTC provides an unconditional monthly payments of $250 or $300 per child. Eligible families may now receive up to $3,600 per child under the age of 6 and $3,000 for those ages 6 to 17. That's up from $2,000 per child. The tax credit system works automatically without the need to signing up at registries had the household filed tax returns for 2019 or 2020, or signed up to receive a stimulus check from the Internal Revenue Service(IRS).

An important and immediate question to ask is whether the CTC program improves the poverty conditions of children. Families receive benefits from the program in form of money and it's important to know if the money is actually being spent for the well being of children. The Center on Budget and Policy Priorities predicts that this policy change under the American Rescue plan will help reduce child poverty by 40%. There have also been reports that parents will leave the labor force after the implementation of the policy. That would hamper the effect of the policy directly on child poverty as parents would lose income and the benefit would only be sufficient to take them back to the previous level. Some policymakers claim this policy change should be made permanent, but it is vital to measure the impact it has on child poverty before taking such a decision. To address these concerns we wish to examine the labor supply impact of the CTC expansion. In particular, we follow the standard literature of labor economics, we expect the tax reform to have impact on intensive margins and extensive margins. The intensive margin is related to the level of work the individual decides to supply and the extensive margin is related with whether or not an individual participates in the labor market. In terms of measurement, extensive margin is typically measured by the number of individuals in paid employment and the intensive margin is measured by the average number of working hours [8].Another takeaway from labor and tax literature is that labor supply is highly responsive margin on the extensive margin to tax reforms [16], therefore this will be our key focal of analysis. Our intuition/hypothesis is that the tax expansion will not change the labor-force participation for those in higher end of the income distribution, but there will be significant incentives for those on the lower end of the income distribution.

The preliminary econometric results suggest that the overall treatment effect support the hypothesis that the temporary reforms of Child Tax Credit in 2021 lays no significant impact on the labor participation decisions and the hours supplied decisions of married couples. This empirical results lends positive evidence that expansion of transfer programs such as CTC might be a good way of reducing household poverty and promote childcare and development without inducing the parents from withdrawing labor participation. Our empirical findings are consistent with recent studies by Ananat and colleagues [1]. Apart from confirming the hypothesis and producing further evidence, this paper contributes to the literature of labor supply responses to tax credit program reforms by addressing some issues that are yet unanswered by existing studies. Relative to simulation and prediction studies based on program parameters, this paper employed a quasi-experimental framework with survey data and addressed the causal relation between the program reform and household labor supply decisions. Beyond confirming the insignificant effect found by Ananat and colleagues, this paper also sought to address the concerns of delayed response.

The labor supply responses may not be immediate due to various sources of frictions such as contracts and knowledge of the program schedule. Compared to the studies of Ananat and colleagues, this paper uses data in 2022 to study how the labor supply decisions are delayed. In this analysis, we specified an Event-Study model with multiple treatment: we consider the implementation in July 2021 as the first wave and the filing of 2021 taxes in April 2022 as the second wave of treatment. Our results add to the literature in several ways. First we confirmed the immediate labor supply response is insignificant because there are information frictions that prevent timely labor supply adjustment from being made. With extended monthly data in 2022 we showed that the labor supply adjustment is delayed for about 7 to 8 month: we observed there is a significant increase in labor force participation rate in the first quarter of 2022 when individuals started to file their taxes and gain additional information about their tax credit eligibility. In addition, we estimated the event study model with multiple events to decompose the total effect of the policy into a sequence of partial effects. The partial effect analysis showed that the second wave of treatment in 2022 induces higher and more significant response than the first wave of treatment in 2021. This observation raises concerns for policy makers and the administrators of social protection programs over possible ways to reduce the information friction. The last piece of result from this analysis is that we showed the partial effects are additively separable by comparing the estimates of the total effect and partial effects. Therefore, this paper provides a more complete anatomy of the dynamics of labor supply responses.

Beyond the standard reduced-form causal study of the impact policy reform, we are also interested in mechanisms of the decision making process. In particular, we are interested whether or not the married couple in a household makes their labor supply decisions jointly or each member of the household make his or her own labor supply decision strategically. Since the labor supply decisions are binary, we can suitably model this potential strategic interaction between married couple using a structural econometric model of static game with discrete choices. Based on the data, model and the estimates of the structural parameter, it can be suggested that a notable component of strategic interaction is present when both the wife and husband of a household make the decision of whether or not to engage in the formal labor market. The coefficients for both the joint effect and the cross effect interaction terms are positive and significant indicating that there are positive externalities from the other member's participation decision through the resources sharing scheme within the household. As a part of the estimation result, we also showed that the presence of children and associated childcare costs are not entirely a negative element to the utility function of parents: the parameter vectors suggested that the parental altruism has a positive and dominant impact on the utility function over the negative impact brought by childcare costs. Subsequently, we can replicate the aforementioned procedures and conduct the semi-parametric structural estimation separately on two sets of data pre and post the policy change. Once we obtain the parameter vectors for both sets, we can compare the parameters before and after the reform. This comparison of the temporal changes allows us to assess the influence of the policy on the strategic interaction between household members.

CHAPTER 2

POLICY AND THEORETICAL BACKGROUND

2.1 Policy Context of Child Tax Credit

Combating poverty has been an essential task for governments across the world for decades. This dedication to treating poverty is regardless of economic systems and ideologies. The origins of social protection programs in the United States can be traced back to the New Deal era of the 1930s and the social safety net expanded through the passing of the Economic Opportunity Act under Johnson's administration [60]. The social safety net was woven to provide assistance to those who are elderly, disabled, or living in poverty for their basic needs such as food, shelter, and healthcare are accessible to all. Amongst various kinds of social protection programs, tax credit programs remain as an important pillar. Tax credit programs provide income support to low-income individuals and families by reducing their tax liability or providing tax refunds. Earned Income Tax Credit(EITC) is the largest in the class of tax credit programs with recipient population over 25 million [35]. Other important tax credit programs include the Child Tax Credit (CTC) program which aim to reduce the financial burden on families with children and support their overall well-being. Child and Dependent Care Credit(CDCC) is similar to the purpose of CTC, but it is smaller in size.

Tax credit programs aimed to families with children are important. It provides the financial assistance for low-income families to send their children to school and this forms an investment in the human capital which is immensely important in the long run [12]. In the meanwhile, tax credit programs are de-signed to preserve the incentives to work. Therefore, such programs help the society to strike a good balance between present period payoff(through work incentives) and future period payoff(through investments in children's human capital). Thus, we will focus on the the Child Tax Credit program and its impact on eligible households and the society as a whole.

2.2 Historic Development of CTC and its Provisions

The Child Tax Credit (CTC) program is a government-sponsored initiative aimed at supporting families with children. The program provides a tax credit to eligible families to help offset the cost of raising children. The credit was intended to provide direct monetary assistance to families with children and to encourage them to have more children.

The CTC has evolved over time to become a significant component of the U.S. tax code together with other credit programs such as the Earned Income Tax Credit(EITC). This section will briefly examine the background and historical development of the CTC program, its eligibility requirements, and the benefits it offers.

The CTC was first introduced in 1997 as part of the Taxpayer Relief Act. The original credit was set at $400 per child, and it was available to families with children under the age of 17. Since its inception, the CTC program has undergone several considerable developments and changes. In 2001, the Economic Growth and Tax Relief Reconciliation Act (EGTRRA) increased the maximum credit to $1,000 per qualifying child and made it partially refundable for families with little or no federal income tax liability. Further expansions were made in subsequent years. In 2017, the Tax Cuts and Jobs Act (TCJA) increased the maximum credit per qualifying child to $2,000. The largest and most significant expansion of the benefit of CTC is under the American Rescue Plan passed in the midst of Covid-19. This expansion will be the main subject of our policy analysis and the details of this expansion will be discussed in the next section. Before getting into the details of the expansion, we will consider how CTC eligibility has changed over time.

Broadly, there are 6 criteria that underly one's eligibility to CTC, but only 2 out of these 6 are of our major interests. To be eligible, the child must be of an age below 17 and this child must either be an U.S national or resident alien. Furthermore, the child must be filed as a dependent. The child must be your son, daughter, stepchild, foster child, brother, sister, stepbrother, stepsister, or any descendent of these individuals. Most importantly, there are income requirements. In addition to the changes in the benefits of CTC program, its income eligibility criteria also changed under these acts. Before TCJA in 2017 the credit began to phase out for taxpayers with income of $75,000 or more for single filers, $110,000 or more for married filing jointly. The credit was completely phased out for taxpayers with income of $95,000 or higher for single filers, $130,000 or higher for married filing jointly. TCJA raised the income thresholds for the phase-out to $200,000 for single filers and $400,000 for joint filers.

Before the CTC expansion in July 2021 under the American Rescue Plan, the program has a "phase-in rate" about 15 cents per dollar of labor earning [17]. The phase-in mechanism entails that it is possible for families to be denied with full benefit provisions because their family income is too low. In particular, researchers have noticed that single parents, ethnic minorities, and families in rural areas are less likely to receive the full benefits [14].

2.3 Child Tax Credit Under the American Rescue Plan

The Child Tax Credit under the American Rescue Plan enacted by the Biden administration from July 2021, is the largest child tax credit ever. The federal government is confident that this tax credit will help all working families succeed. There are some major changes under the ARP. First, the income testing threshold was made higher: families with a combined income of up to $150,000 or a single parent income of up to $112, 500 would receive a maximum of $3000 per qualifying child of age 6 to 17. The maximum benefits are set to $3600 for children below 6 years old. Secondly, the tax credit increased the age limit for children eligible to receive benefits from 16 to 17. Thirdly, ARP made CTC fully refundable, which means that eligible families can receive the full amount of the credit even if they don't have any federal income tax liability. In addition to the making the CTC refundable, the ARP removed the phase in mechanism with the aim that more eligible families can receive the full amount of benefits. Finally, the benefits can be claimed through monthly installments: eligible households can receive a maximum monthly transfer of $250 per qualifying child and $300 per qualifying child under 6 years old. These features jointly suggest that under the ARP, the CTC program provides a near universal transfer to eligible families.

The phase-out mechanism follows a two-staged process: Initially, for single parents earning more than $112,500 and married couples earning more than $150,000, the credit gradually decreases by 5% of their adjusted gross income until it returns to the levels seen before 2021. Secondly, the credit's worth is additionally lowered by 5% of their adjusted gross income exceeding $200,000 for single parents and $400,000 for married couples. Having these parameters clarified, we can now visualize the level of benefits and the (adjusted gross) income.

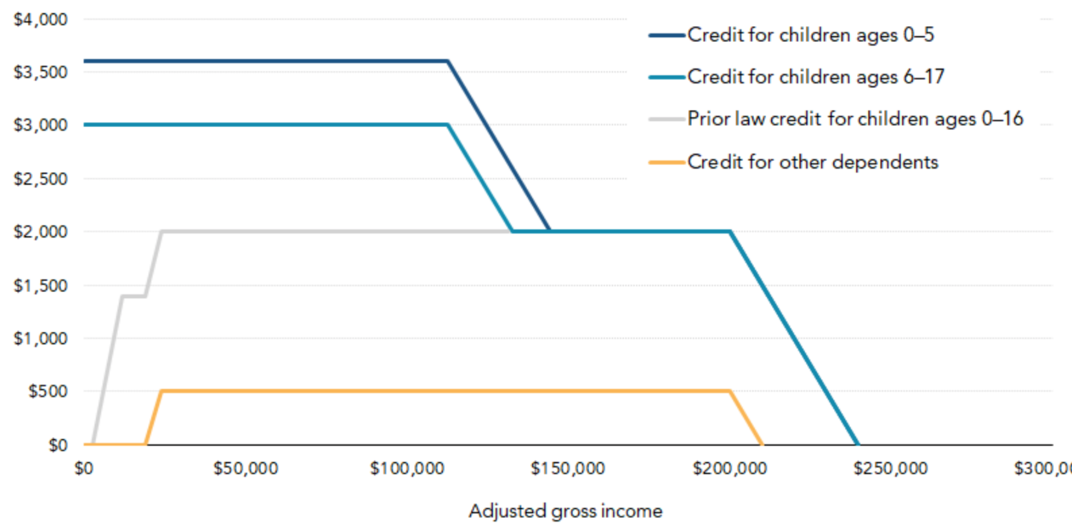

Figure 2.1: The figure gives the relationship between the level of benefits and the adjusted gross income(AGI) assuming that the single filer has exactly 1 eligible child.
Source: Urban Institutes & Brookings Institution Tax Policy Center

It can be seen that the ARP removes the phase-in mechanism as the benefit-AGI curve is horizontal from $0 up until the AGI for the first phase-out threshold. In comparison, the grey curve(pre-ARP parameters) slopes positively from $0 suggesting that the phase-in mechanism does exclude extreme low income earners from claiming the full benefits. Another observation is that the first phase-out income threshold occurs at a lower income level relative to the single phase-out income threshold stipulated by the TCJA, but the benefit reduction rate is not changed. Combining these observations, we expect that under the parameters of ARP, eligible householders can better receive payment and thus adjust their budgets and labor supply decisions accordingly.

Apart from the changes in program parameters, several operational aspects were also improved. Not only expanding the benefits and eligibility threshold, but the policy change also streamlined the benefit receiving process as all families filing tax returns would receive benefits through direct deposit or mail, and families not filing tax returns being eligible to enroll into the tax credit through a nonprofit, Code for America's campaign using mobile phone application. Together with this new refundable credit, advanced payments were made available under ARPA. Eligible families can receive advance payments of up to 50% of their estimated credit.

2.4 Theoretical Intuition of the Labor Supply Response

From the structure of the CTC program, it is obvious that eligible agents will face a different set of budget constraints. The pre-2021 parameters will create a "hump-shaped" budget constraint just like the case of EITC, the ARP changed the CTC by removing the phase-in mechanism. Thus, we would expect that before the households' income exceeds the first phase-out threshold, the CTC benefits act like a lump-sum non-labor income for eligible households. The change in the shape of budget constraints will induce a combination of behavior responses to changes in the implied incentives. We will graph out three generic budget constraints: the original budget constraint without any credit programs at all; the "hump-shaped" budget constraint with the pre-2021 parameters; the budget constraint with the ARP parameters.

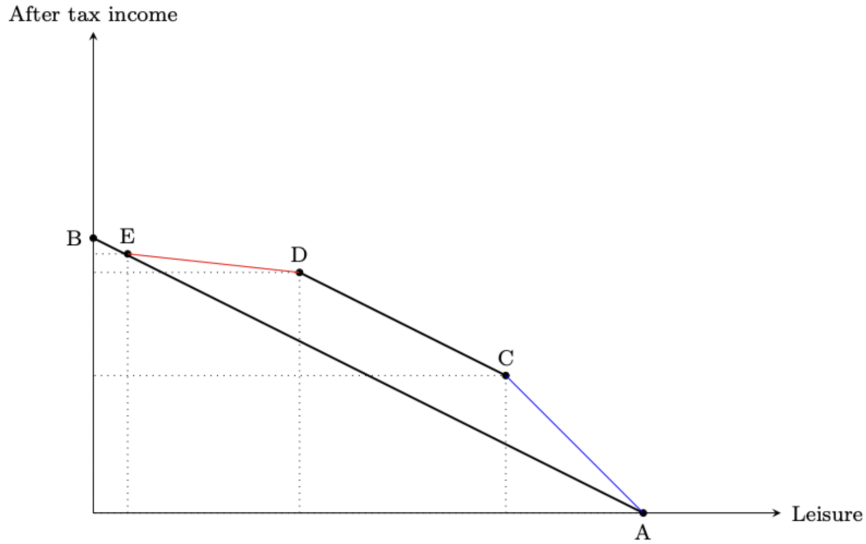

Figure 2.2: The slopes and intercepts are generic and for demonstrative purposes, they do not correspond exactly to the policy; This budget constrain correspond to the CTC policy before the ARP parameters and is similar to EITC

The budget constraint is graphed on the after tax income-leisure plane. The solid line AB represents a typical individual's budget constraint without any social protection programs. He will choose how much leisure to consume and from this he also makes a decision regarding his labor supply since the time endowment is fixed. The colored line ACDEB represents the simplified budget constraint that an eligible CTC recipient faces with the pre-2021 parameters. This budget constraint is drawn for illustrative purposes and the slopes do not correspond in magnitudes to the program parameters. In addition, we graphed the budget constraint assuming that the phase-in period follows a one stage process for graphing conveniences.

The introduction of the CTC program with pre-2021 parameter improves the well-being of the tax-payer since for every choice of hour supplied, the available leisure-income bundle is higher than the original case. Thus, this observation suggests that the introduction of CTC provides no negative incentive to drop labor participation for those who already participate. It is plausible that the additional income support provided by the program may incentivize those who do not yet participate to change their decisions.

The impact of the introduction of CTC on intensive margin for those who already participate should be considered by a case-by-case approach: the substitution and income effect may vary depending on which part of the budget constraint the tax-payer is on. The wage subsidy carries both substitution and income effect for tax-payers on the section under the blue line AC. In this region, the substitution effect is positive and incentivizes tax-payers to supply more hours and the income effect is negative. Thus, the net effect on hours supplied for tax-payers in this region is not clear. For the region under the black line CD, the pre-2021 CTC parameter suggests that only income effect is present and it has a negative impact on hours supplied. In the region under the red line DE, the impact on hours supplied is negative: similar to previous two regions, the income effect is also negative on hours supplies; the benefit reduction and phase-out mechanism in this region implies that the substitution effect is also negative on hours supplied. Therefore, the combined behavioral response in hours supplied to the program on the red region is negative in net. The final possibility for the intensive margin is that those who original on line EB may segment switch by reducing their hours supplied and claim the credits.

We can graph the budget constraint of a tax-payer under the ARP parameters. Following the rules of ARP parameters, the program becomes an universal basic income scheme with two phase-out periods. The first phase-out is to the pre-2021 level and then the second phase-out is the same as pre-2021 phase-out. The only changes to the budget constraint is on the region from point A to point C and some parts on the line CD.

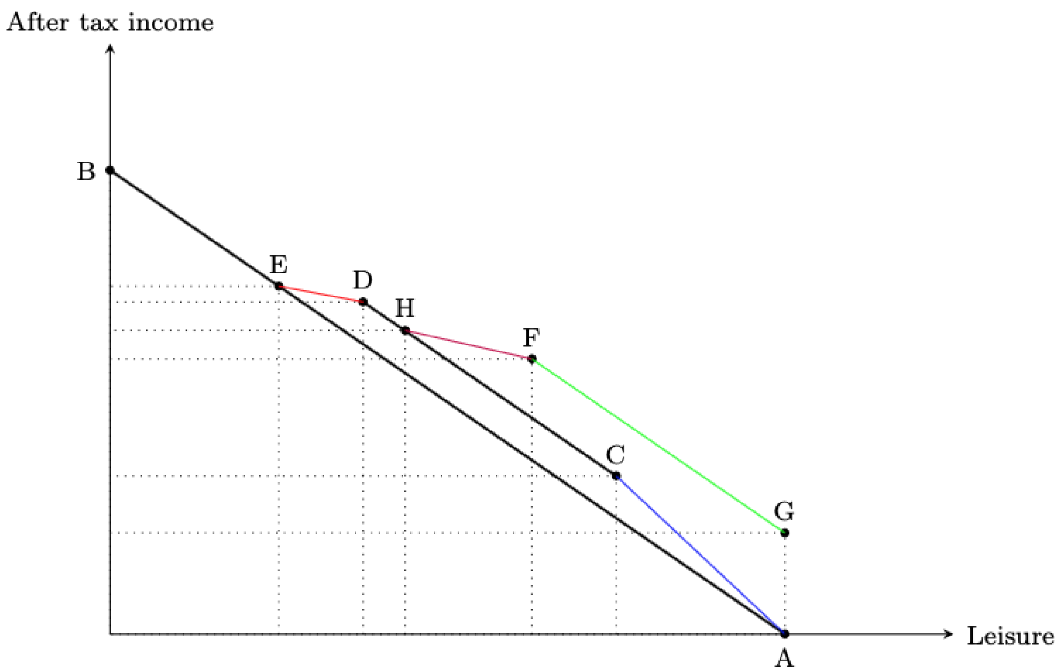

Figure 2.3: The slopes and intercepts are generic and for demonstrative purposes, they do not correspond exactly to the policy.

The budget constraint under the ARP CTC parameters is the colored line GFHDEB. It should be noted that the budget constraint to the left of point H is exactly identical to the budget constraint with the pre-2021 rule, thus we should not expect any differences in behavioral responses in labor supply decisions to the left of point H. Compared to the pre-2021 rule, it is obvious that the budget constraint with the ARP parameters is higher than or equal to budget constraint with pre-2021 rules, thus we can infer that the removal of the phase-in mechanism and expanded benefits will have a non-negative impact on labor participation decisions relative to the pre-2021 rules. In particular, the elimination of the phase-in and expanded benefits can address the fixed cost of participation for more credit-constrained parents [23]. Indeed, the CTC benefits have been used by recipients to fund childcare expenditures [51].

The conditional hours supplied should also be discussed in a case-by-case fashion since the behavioral response is dependent on the segments of the budget constraint. Since the region under line HDEB is also the budget constraint with the pre-2021 rules, the behavioral would be the same: income effect has a negative impact on hours supplied under the region below HD; the region under DE has an unambiguous negative impact on hours supplied; there is a possible segment switching behavior on the region under EB which may also negatively affect the hours supplied. We would expect a change in the behavioral response to the right of point H. The universal basic income nature of Child Tax Credit under the ARP regime shifts the original black budget constraint parallelly up to line GF before the first phase-out begins. Relative to the pre-2021 and no tax credit budget constraint, the ARP parameters induce a negative income effect on the hours supplied on the region below line GF. The line FH is the first phase-out back to the constant benefit level of the pre-2021 rule. The intensive margin effect on this region is unambiguous: both substitution and income effect is negative on the hours supplied. Therefore, combining these cases, the overall impact on the hours supplied depends on the income and we should incorporate this consideration into the reduced-form estimation.

2.5 Preliminary findings on the labor supply impact of tax credit programs

Main stream prior research on tax credit program has been concentrated on the Earned Income Tax Credit since it is amongst one of the largest US social protection programs. A large branch of empirical studies on EITC focus on the labor supply decisions of single mothers. Eissa and Liebman found that 1987 expansion of the EITC produced positive labor participation responses and they admitted that the estimates on hours supplied are less robust [24]. Meyer and Rosenbaum confirms the findings of Eissa and Liebman using a more complex set of models and set up a structural model which confirms that income and leisure are perfect substitutes [46].Raj Chetty and Emmanuel Saez used an experimental approach to study why the response in extensive margin is more salient than the intensive margin responses [13]. Others have also examined the impact on married couples since they face a different incentive design than single parents. Eissa and Hoynes studies the extensive margin responses to EITC expansions and they found differing behavioral responses between the husband and wife in the household [22].

Compared with the relevant literature of EITC, the studies of Chilf Tax Credit program are limited because the program parameters of CTC is similar to EITC and it is difficult to separately identify the CTC impact from EITC impact [44]. Therefore, the main stream research works on CTC focus on sources of identification prior to 2021 expansion. A popular candidate for exogenous variation used in empirical studies of CTC is the loss of CTC when children/dependent reaches the age of 17 [26]. Feldman and colleagues found that the implied intensive margin elasticity is about 0.3 and this indicates that the intensive margin response is not significant. Similarly, the loss of dependent status as children age has also been used as a source of exogenous variation [41]. They found a higher intensive margin elasticity of 0.75 compared Feldman et.al's result. The discontinuities in child age at the end of tax year can also be harnessed to identify the effect of Child Tax Credit on labor market responses [61]. The authors found that the discontinuities between December and January and the implied differences in benefits have a negative effect on the maternal labor supply decisions.

CHAPTER 3

DATA

This paper will investigate the causal impact of the CTC expansion on child poverty conditions and labor supply of households. To address the labor-force participation decision, we need data at the resolution at least at the households levels. Therefore, a plausible candidate is the Current Population Survey(CPS). CPS is a monthly survey conducted by the U.S. Census Bureau and the Bureau of Labor Statistics to collect data on labor force characteristics, such as employment status, occupation, industry, hours worked, and earnings, as well as demographic characteristics, such as age, sex, race, educational attainment, and family status. The survey is administered to a nationally representative sample of households and individuals, and has been conducted since 1940. The CPS data is also rich in its coverage of socio-economic indicators as well as labor market indicators. Since the CPS data is recorded at household/individual levels, we believe this is the preferred data source to use for our research design. Since the CTC expansion act was passed on 15th July 2021, to inspire the Diff-in-Diff typed identification we would require at least two waves of data before the policy and one wave of data after the policy to establish the difference pre and post the policy change. Since the policy change is set to be temporary and specific for the year of 2021, we decided to use the CPS data from January 2021 to December 2021. We will use an extended monthly data in 2022 in later sections to discuss the plausibly delayed response and various frictions. This point is discussed in more details in following sections.

3.1 Sample Construction

From the CPS monthly household level data, we focus on individuals between age of 18 to 65 since they remain active in the labor force. The initial pooled data contains 860,724 observations, we form our sample by selecting only the married households and households with a detailed record of their labor-force participation. This preliminary selection leaves us with 439,282 observations. Each individual is identified by a household id, this is stored as *serial* in the CPS data. This forms our baseline sample of analysis, but we also form two sub-sample by the recorded sex: we call it the husband sub-sample and the wife sub-sample. We did this because we hypothesize that there are heterogeneous reactions to CTC policies based on gender differences and the related financial incentives.

3.2 Measurement and variable Selection

Given the richness of the CPS data, we are able to analyze the labor market responses of individuals both on the extensive margin and the intensive margin. For the analysis on the intensive margin, we are interested in examining how the weekly hours worked change as the result of the policy change. For the extensive margin analysis, we are interested in the changes in labor force participation decisions. For the labor force participation, we use variable *labforce* to measure binary participation. For the weekly hours, we use variable *ahrsworkt* to measure the total number of hours the respondent was at work during the previous week. For employees, this refers to the duration of their work hours. For family members who are not compensated for their work, this includes the amount of time spent doing tasks related to the family business or farm, but not housework. For employers and self-employed individuals this reflect the time they spent on managing their businesses.

The labor supply/participation decisions of individuals are made by considering various socio-economic conditions. To account for this and to get consistent estimators, we wish to control for these demographic variables in the reduced-form estimations. To be consistent with the literature, we include common demographic variables such as age, age squared, sex, race, level of education, family income. The family income is denoted by a categorical variable *faminc* in CPS data. This encompasses all types of monetary earnings received by an individual, including income from employment, net profits from businesses or farms, rental income, pension payments, dividends, interest, social security benefits, and any other sources of money income. The level of education is also characterized by a categorical variable *educ*: this ranges from no schooling at all to doctoral degrees. Relevant to our policy analytical context, we include the number of children, the age of youngest children, and the age of the oldest children. These information, together with the family income, inform us the household's eligibility for the CTC program. Other relevant control variables we include are the wave of the data, the state of the household. Finally, we also include the variable *hwtfinl* to capture the household weights which we later use in the econometric analysis. Due to the complexity of the stratified sampling technique used by the CPS data, we will not discuss the details of the algorithm underlie the household weights. Interested readers can refer to the Current Population Survey Technical Documentation [10] for more elaborated discussions.

3.3 Descriptive Statistics and data visualization

We can first produce the summary statistics of our variables by levels of education. We will create three groups: those with education attainments less than high school, those with high school education, and those with education attainments above high school. The following table presents the weighted means and standard deviations by education levels:

Table 3.1: Weighted Summary Statistics of selected variables for the pooled sample

	(Less than High School)		(High School Education)		(Education above High School)	
	mean	sd	mean	sd	mean	sd
Family Income	53266.85	42249.12	77817.36	51905.81	122244.14	61788.92
White	0.82	0.38	0.82	0.39	0.80	0.40
Female	0.47	0.50	0.48	0.50	0.54	0.50
Age	46.86	11.35	47.29	11.95	45.87	11.19
Number of Own Children in Household	1.69	1.52	1.14	1.27	1.19	1.19
Labor Force Participation	0.62	0.49	0.71	0.45	0.81	0.39
Observations	25684		97649		265303	

In calculating the mean of family income, we assume the household earns exactly the median income within the income interval for the reported category. Female and White can be interpreted as the percentage of Female/White in the sample. It can be seen that ethnicity and the number of own children are distributed evenly across different education levels. Similar description applies to gender, but we observe that females generally are more likely to have education attainments above high school. In particular, we note that family income has a apparent that family income increases with the level of education. Similarly, people with more years of education are more likely to participate in labor market.

For the outcome variables of our interests, we need more thorough examination than means and standard deviations. We can use a two sample test to determine the differences between groups. Instead of partitioning the pooled sample by education, we wish to separate the sample by CTC parameters.

Table 3.2: Summary statistics on key socio-economic and demographic variables by treatment status

	(Treated)		(Control)		(Difference)	
	mean	sd	mean	sd	Difference in Means	t-statistics
Family Income	75512.91	37230.34	122506.26	67630.39	44057.47***	(279.96)
White	0.80	0.40	0.81	0.39	0.01***	(6.99)
Female	0.51	0.50	0.52	0.50	0.02***	(9.82)
Age	40.05	8.55	49.59	11.35	9.98***	(331.31)
Labor Force Participation	0.79	0.41	0.76	0.42	-0.03***	(-26.17)
Weekly Hours Supplied(conditional on participating)	39.82	11.64	40.58	11.88	0.66***	(14.71)
Observations	133698		254938		439282	

We define households with children and family income below $150,000 to be the treated group and the rest of the sample to be the control group. By construction, the treated group have access to CTC benefits above pre-2021 level and control group is constituted by householders without access to the expanded provisions. For the outcome variables of interests, there are significant differences between two groups: the group eligible to expanded CTC benefits have higher labor participation rate but lower weekly hours supplied conditional on participating. From the time series plots, we can confirm that such difference persists within every wave of data.

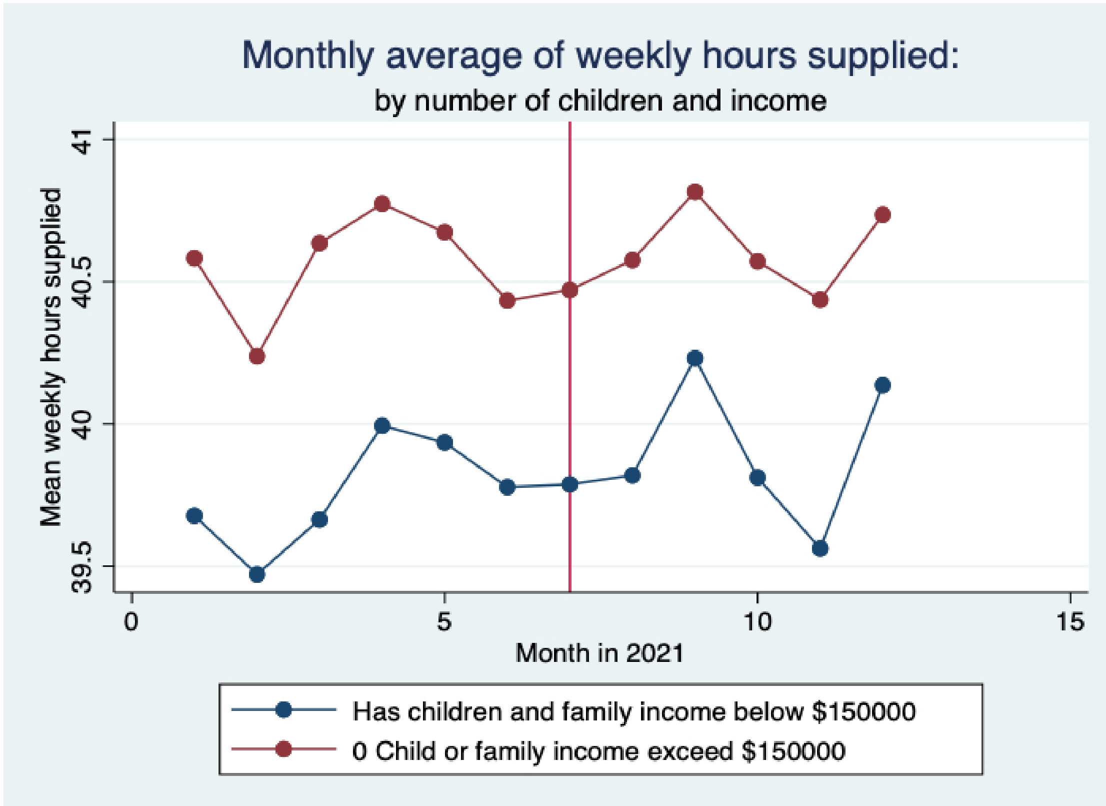

Figure 3.1: The scatter is the monthly mean of weekly hours supplied and we simply connect these scatters over all months in 2021

The vertical distance between two lines shows that there exists marked differences between two groups' weekly hours supply over the span of our data. For all waves of the data, the control group has higher weekly hours than the treated group. This can be attributed to the intrinsic demographic differences underlie two groups. Despite such differences, it is obvious that two groups' weekly hours supplied follow very similar trends:both groups experienced a dip in weekly hours in the first two months of 2021 and the gradually recover towards the third and fourth quarter of 2021. Consistent with the two sample test, the treated population has higher binary participation than the control group for all waves data.

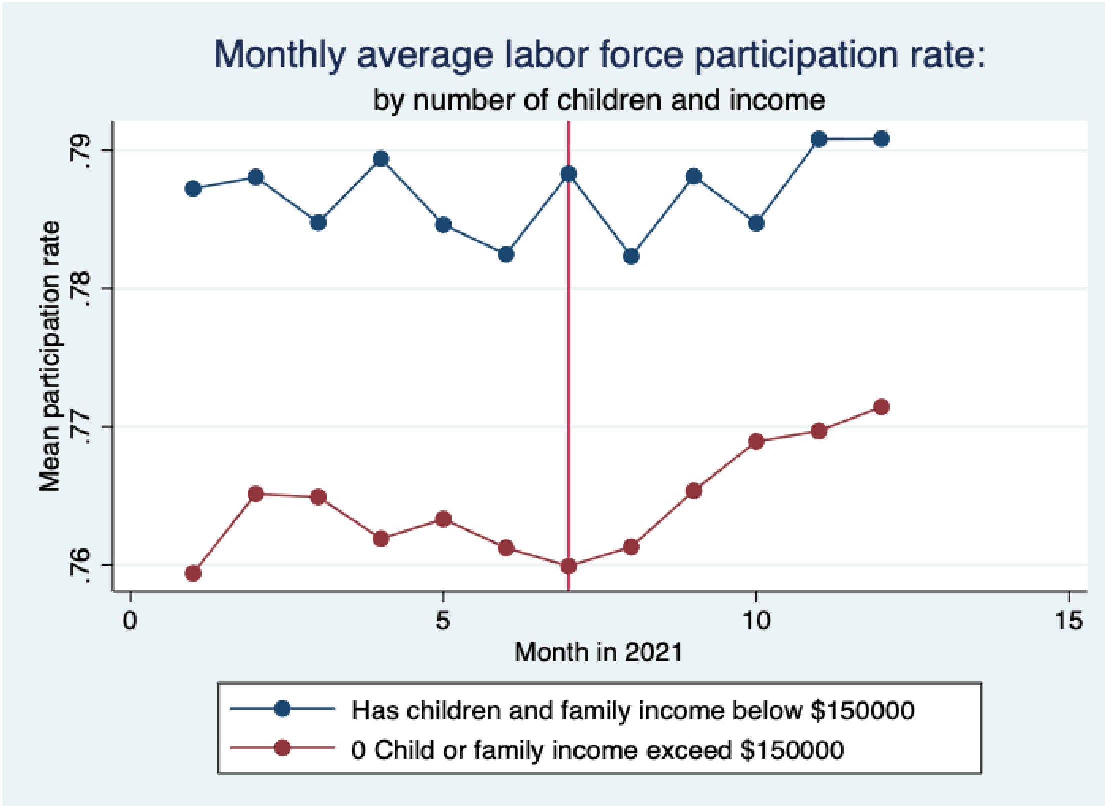

Figure 3.2: The scatter is the monthly mean of labor participation and we simply connect these scatters over all months in 2021; the red vertical line is the month(July) when CTC was expanded

Trends of two groups are not perfectly parallel as the case for weekly hours supplied. However upon crude visual examination, one can observe that the trends are grossly similar: binary participation generally was declining prior to July and the participation after July gradually recover for both groups.

CHAPTER 4
EMPIRICAL METHODOLOGY

By the particular feature of the transfer program, we can form a natural experiment design to study the causal impact of the CTC expansion of extensive margins of household labor supply decisions. Since the transfer is made on per child basis, then we claim that households without any children provides a good comparison for those families with one or more children. The following figure gives the distribution of the number of children of the selected households in our sample.

The validity of this quasi-experimental approach relies on the parallel pretrend assumption: for the causal relation to be identified by our strategy, we need to establish that there is no differing underlying trends in labor-force participation between our control group(no children) and treatment group(more than 1 child). Another requirement for our identification strategy is that there was no contemporaneous shocks to labor markets other than the CTC expansion [7]. A particular concern we will evaluate in the robustness checks is the no contemporaneous shocks assumption: the CTC expansion was used as a policy instrument on the US government side to stimulate the waning economic activity caused by the coronavirus, there are valid reasons to believe that the labor market condition was subject to multi-dimensional shocks throughout the period of our study. Standard theory predicts individuals of different socioeconomic characteristics have different incentives and thus different responses to the tax credit expansion, thus we also aspire to estimate the heterogeneous response by estimating the model using sub-sample. The DID approach is based on strong identification assumptions such as the parallel trend assumption:in absence of the treatment, the average outcomes for the treated and control groups would have followed parallel paths over time. This assumption may fail to hold in various circumstance and the DID estimation of the treatment effect can be biased. We should consider how these threats may materialize in the particular case we study in the robustness checks. Before delving into such discourse, we must explicitly define what the theoretical estimand of our empirical methodology and their interpretation.

4.1 Theoretical Estimands, empirical estimates, and Identification

We can use the language of the potential outcomes to explain our empirical methodology [54]. Let Y_i be the outcome variable of interests and let t_i be the binary treatment variable such that $t = 1$ if the household i is treated 0 if not. Then in our case $t = 1$ if the household is deemed to be eligible to the expanded CTC benefits. This designation is consistent with relevant literature of tax credit programs [24]. We let $Y^1_{i,t}$ denote the potential outcome of individual i at time t had the household i been treated. Similarly, $Y^0_{i,t}$ denote the potential outcome of individual i at time t had the household i not been treated. Only one of these two is observable and the other is called the counter-factual. Then we can define the individual level treatment effect to be difference between the observed outcome and the potential counter-factual outcome: $Y^1_{i,t} - Y^0_{i,t}$.

Then what we would like to estimate is the Average Treatment Effect such that:

$$ATE = \Delta = \mathbb{E}(Y^1_i - Y^0_i) \tag{4.1}$$

Therefore, our unit specific quantity is the difference in labor force participation of individual i if he/she had been eligible for CTC program and the target population of units is thus married couples in the US. However, recognizing the inherent differences in socio-economic and demographic traits between the treatment group and control groups analyzed in previous section, the empirical estimation using the experimental potential outcomes is biased since the treatment status is not fully randomized.

To get an unbiased estimator of the average treatment effect, we use a quasi-experimental identification, Difference in Difference(DID) to allow for the inherent differences between control and treatment groups. This identification strategy has a looser requirement: this requires the control group and the treatment group to have the same pre-trend of labor-force participation before the treatment is imposed. Given our research design and the identification strategy, we estimate only the Average Treatment Effect of the Treated(ATT) such that:

$$ATT = \mathbb{E}(Y_i^1 - Y_i^0 | t_i = 1). \tag{4.2}$$

4.2 Identification Assumption

We shall focus on the identification assumptions for the proposed DID estimator to have a causal interpretation. The following properties are necessary for any estimators to have a causal interpretation. The first assumption we make is the exchangeability assumption. This requires that the distribution of potential outcomes is independent of the treatment assignment design mechanism [34]. The second assumption/requirement is the Stable Unit Treatment Value Assumption(SUTVA).This assumption is necessary for every estimator to have a causal interpretation [53]. First part of this assumption requires the units to be well defined and the designated treatment status remains stable in our data. The second part of this assumption requires that treatment effect does not spillover: we require the treatment status of one unit does not affect the potential outcome of another unit. The third assumption is the positivity assumption. This requires that conditional probability of receiving the treatment is strictly greater than 0 and strictly less than 1 [34].

Pertaining to our Diff-in-Diff and Event Study empirical design, we need the following assumptions to hold to get consistent estimator. The first and perhaps most important assumption is the parallel pre-trend. We require that the outcome variable for the control and treatment group to trend similarly before the policy shock. In fact, the most ideal situation is to have parallel post trend: control population and treatment population would have trended similarly had the treatment not imposed. But this is counter-factual and cannot be tested. Thus, we can only use the observable parallel pre-trend to approximate the parallel post trend. In justifying this assumption, we produced visual examinations and a series of statistical tests. There are several ways in which this assumption may fail in real world policy evaluation. Consider the Ashenfelter's Dip for training programs: the sharp fall in earnings of the treatment population can introduce a bias in the estimated treatment effect because what we estimate using DID is essentially a mixed effect of training program and catching up effects [31]. Another way this assumption fails is through the anticipation effect. Angrist and Pischke noted that if individuals can anticipate the policy change, they may adjust their behavior before the policy change and this can cause violations to the common trend assumption [2]. A further issue regarding the pre-trend testing is that it can be of low statistical power [52]. This implies that failure to reject the parallel pre-trend assumption might occur, but this failure to reject does not imply the parallel post trend which more rigorously define the counterfactual. Therefore, there is still a potential possibility of bias in the Diff-In-Diff estimator even if parallel trends test holds.

The second key identification assumption is the common shock assumption. This requires that there ought to be no contemporaneous shocks that affect outcome variables of both groups differentially over the period. This means that unobserved shocks move the outcome variables of two groups in the same direction. Any violations to the common shock assumption can introduce bias in the estimates. This assumption is particularly important because over the year of 2021 there was a number of policies enacted that might affect labor supply decisions. This assumption is difficult to be justified by statistical tests, a good practice would be to examine the sensitivity of the results to alternative models [33]. We shall examines how these assumptions may be violated for our particular design in the following sections.

4.3 Potential Threats to Identification Assumptions

There are several ways in which the parallel pre-trend assumption fails to hold in our case. One potential violation can be the presence of the cohort effect. In the summary and descriptive statistics we can see that the age variable differs violently between the control and treatment groups. This can raise concerns of the cohort effect where two groups are in different points of their life and there might be a cohort-specific labor supply decisions that can invalidate our identification assumption [37]. A second potential threat to the parallel pre-trend is heterogeneous response to business cycles and this relates to the existence of contemporaneous shocks to households [21]. The CTC program is only one of many items in the collection of policies enacted by Biden's government in the post-Covid era, and it should be suspected that other policies do affect the labor-force participation thus constituting a contemporaneous shock. However, we argue that despite the presence of these contemporaneous shocks, our identification can still be valid and bare a causal interpretation as long as those contemporaneous shocks affect the labor-force participation of in similar fashions. That is as long as the contemporaneous shocks do not cause heterogeneous response in labor participation between the control and treatment group, the pre-trends are shifted vertically by these shocks, but they remain parallel. Beyond a visual scrutiny of the time series plot, we should address this typical concern with greater details in following sections.

CHAPTER 5

BASELINE ECONOMETRIC RESULTS

In this section we present the estimation results from our baseline DID estimates using the OLS method. We are concerned about the heterogeneous treatment effect over different sex, thus we estimate the same set of specification for the pooled population and for the male sub-sample as well as the female sub-sample. For each (sub)sample, we estimate four specifications and have the results presented in separate columns.

The first specification we estimate is:

$$y_{ist} = a + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 Treat_i Post_t + \epsilon_{it}. \tag{5.1}$$

This is the simple DID with no additional control variables and the average treatment effect is captured by parameter β_3.

The second specification we estimate is the same as 5.1, but we add a rich set of control variable:

$$y_{ist} = a + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 Treat_i Post_t + X_{ist} + \epsilon_{it}. \tag{5.2}$$

where X_{it} is a vector of co-variates including age, squared age, sex, race, and number of children. This is the baseline results for our causal analysis for two reasons. First, controlling for these co-variates can serve to absorb potential confounding variations that poses threats to our identification assumptions. Second, The within-group variations in co-variates can help to reduce the error variance of the regression. Thus, the statistical tests and inferences drawn from the tests will be of higher statistical power.

Alone the same line of thinking, we further control for state fixed effects and the impact of education in 5.3 and 5.4 respectively:

$$y_{ist} = a + \delta_s + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 Treat_i Post_t + X_{ist} + \epsilon_{it}, \tag{5.3}$$

$$y_{ist} = a + \delta_s + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 Treat_i Post_t + Educ_{ist} + X_{ist} + \epsilon_{it}. \tag{5.4}$$

The following table shows results for baseline's pooled sample regression outputs:

Table 5.1: Estimated results for the pooled sample

	(1)	(2)	(3)	(4)
VARIABLES	DID	DID: Regression Adjustment	DID: State FE	DID: State FE and education adjusted
treat_eff	0.00386	0.00149	0.00129	0.00262
	(0.00282)	(0.00262)	(0.00262)	(0.00261)
age		0.0459***	0.0459***	0.0376***
		(0.000135)	(0.000134)	(0.000140)
sex		-0.0896***	-0.0895***	-0.0956***
		(0.000875)	(0.000874)	(0.000861)
race		-3.44e-05***	-2.70e-05***	-4.39e-05***
		(2.78e-06)	(2.89e-06)	(2.84e-06)
nchild		-0.000675	-0.000882	0.00344***
		(0.000569)	(0.000570)	(0.000569)
Constant	0.566***	-0.392***	-0.425***	-0.238***
	(0.00120)	(0.00420)	(0.00495)	(0.0103)
Observations	1,116,407	1,116,407	1,116,407	1,116,407
R-squared	0.030	0.303	0.304	0.327
State FE	NO	NO	YES	YES
Regression Adjustment	No	YES	YES	YES
Education Control	No	No	No	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 displays the result for regressing labor supply without any control variables. The treatment effect is very small as we can see a 0.03 percentage points increase in labor force participation as an effect of Child Tax Credit. This estimated average treatment effect is positive and insignificant. Since the specification of Column 1 is done without the controls, the treatment's impact is likely to be biased. Column 2 introduces control variables for age, sex, race and number of children. When controlled for those demographics, the statistical significance of treatment effect changed, it is insignificant suggesting around 0.15 percentage points increase in labor force participation for the treatment group as an effect of the Child tax Credit. Adding state-fixed effects (Column 3) and educational level controls (Column 4), the treatment effect does not change drastically, it is positive and insignificant still. These results suggest that the expansion of the CTC provision causes about from 0.1 to 0.3 percentage point in labor participation of the pooled population. In all scenarios, we do not see any significant negative effect of the Child Tax Credit expansion on the labor supply of the recipient families. The standard errors for the specifications do not change across columns, as they remain steady at 0.2 percentage points. The standard errors are not large enough when compared to the coefficients hence justifying the insignificance of the results.

These findings are consistent with the observed trends in mean labor-force participation rate of control and treatment groups displayed in trend plots in the previous sections as well as findings in the relevant labor supply and CTC literature [1].

This preliminary result is a weighted averaged of the husband and wife specific effect. Therefore, it is of our analytical interest to split the pooled sample and estimate our reduced form models again for the husband and wife sub-samples respectively. We estimate the same set of equations for the the husband of the household and the following table presents the results:

Table 5.2: Estimated results for the male sub-sample

	(1)	(2)	(3)	(4)
VARIABLES	DID	DID: Regression Adjustment	DID: State FE	DID: State FE and education adjusted
treat_eff	0.000427	-0.00171	-0.00190	-0.00112
	(0.00326)	(0.00305)	(0.00305)	(0.00305)
age		0.0502***	0.0503***	0.0426***
		(0.000190)	(0.000190)	(0.000198)
race		-3.77e-05***	-3.27e-05***	-5.01e-05***
		(3.69e-06)	(3.82e-06)	(3.76e-06)
nchild		0.0223***	0.0218***	0.0241***
		(0.000651)	(0.000652)	(0.000661)
Constant	0.614***	-0.612***	-0.642***	-0.434***
	(0.00168)	(0.00552)	(0.00656)	(0.0144)
Observations	537,075	537,075	537,075	537,075
R-squared	0.053	0.345	0.347	0.368
State FE	NO	NO	YES	YES
Regression Adjustment	No	YES	YES	YES
Education Control	No	No	No	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5.2 above, shows the regression output for the male sub-sample in the data. This table is striking in a sense that the treatment effect for all regressions(Column 1-4), are small in magnitude and not statistically significantly different from zero. One should note that depending on the specification, the estimated effect on participation changes in signs. This should not cause any concerns since they are not significant at usual levels. One should not interpret it as the Child Tax credit program has a negative effect in the labor force participation of the male population, in the families with 1 or more children. The standard errors in the table are fairly smaller being stable around 0.03 percentage points across all columns. The results support the stability of our baseline estimate in column 2. This result offers an alternative evidence to the negative work incentive for males found in similar tax credit programs' expansion such as Earned Income Tax Credit(EITC) [24].

Table 5.3 lists the regression outputs for the female sub-sample in the data:

Table 5.3: Estimated results for the female sub-sample

	(1)	(2)	(3)	(4)
VARIABLES	DID	DID: Regression Adjustment	DID: State FE	DID: State FE and education adjusted
treat_eff	0.00660	0.00416	0.00395	0.00550
	(0.00416)	(0.00396)	(0.00395)	(0.00390)
age		0.0417***	0.0417***	0.0327***
		(0.000189)	(0.000189)	(0.000198)
race		-2.96e-05***	-1.98e-05***	-3.34e-05***
		(4.07e-06)	(4.23e-06)	(4.16e-06)
nchild		-0.0204***	-0.0202***	-0.0135***
		(0.000865)	(0.000866)	(0.000860)
Constant	0.519***	-0.434***	-0.472***	-0.325***
	(0.00170)	(0.00551)	(0.00667)	(0.0139)
Observations	579,332	579,332	579,332	579,332
R-squared	0.019	0.262	0.264	0.291
State FE	NO	NO	YES	YES
Regression Adjustment	No	YES	YES	YES
Education Control	No	No	No	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table is different from the ones above. Without the demographic and educational controls and state fixed effects, the treatment effect of the Child Tax credit is positive, insignificant but small in magnitude. The signs of the average treatment effect are the same over different specifications. The coefficient for the treatment effect is stable about 0.1 across four specifications, suggesting an increase in labor force participation by 0.1 percentage points as the effect of Child Tax Credit. The standard errors are similar around 0.3 percentage points across all 4 specifications, but the magnitude of the effect in Column 1, makes the coefficient significant. This suggests that the causal impact of Child Tax Credit for the female sub-sample was positive small, but significant. The direction of effect is different from what is expected theoretically, but the results speak to the empirical findings of some tax credit expansions of similar kind. Eissa and Liebman found a positive impact of EITC on labor-force participation for eligible female(with children) relative to ineligible female(childless) [24]. Similar findings were found by Meyer and Rosenbaum [46].

The practical implication of the above results is encouraging: the expansion of CTC successfully reduce the poverty rate without incurring negative behavioral response of reducing labor-force participation.

5.1 Hours Supplied

The primary focus of this study is on the more robust participation response [13], but we will also present the estimated impact on hours supplied for the completeness of the theme of analysis.

Previous results are suggestive of the fact that expansion in CTC provision does not induce a reduction in labor force participation. To have a complete image of the behavioural response of married couples to this expansion, we need to examine the responses along the intensive margin. To do this, we can estimate the same specifications with another dependent variable: hours supplied. Unlike the previous part where the dependent variable is binary, we have conducted the winsorization at 1 percent level of the hours supplied to check the impact of extreme values on the casual impact of expansion on hours supplied. Furthermore, for the analysis of the intensive margin, we focus on the responses in hours of those who are already employed. This is because we suspect the unemployed population may be subject to different shocks than the employed population and may be of different un-observable characteristics. The quasi-experimental estimator we specified will likely to be biased and inconsistent if we fail to account for these threats to identification. Thus, we will consider the responses in hours supplied conditional on those who are employed.

The following table presents the estimates using the pooled sample:

Table 5.4: Estimated results for hours supplied using the pooled sample

	(1)	(2)	(3)	(4)
VARIABLES	DID	DID: Regression Adjustment	DID: State FE	DID: State FE and education adjusted
treat_eff	2.984	2.787	2.546	2.340
	(2.372)	(2.336)	(2.334)	(2.332)
age		-6.158***	-6.208***	-5.179***
		(0.188)	(0.188)	(0.192)
sex		-11.22***	-11.54***	-9.469***
		(0.817)	(0.816)	(0.822)
race		0.0110***	0.00687***	0.0117***
		(0.00252)	(0.00261)	(0.00262)
nchild		-2.744***	-2.336***	-2.322***
		(0.460)	(0.461)	(0.463)
Constant	155.8***	533.8***	509.2***	478.6***
	(1.028)	(5.353)	(5.759)	(13.03)
Observations	663,419	663,419	663,419	663,419
R-squared	0.001	0.023	0.026	0.028
State FE	NO	NO	YES	YES
Regression Adjustment	No	YES	YES	YES
Education Control	No	No	No	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 of table 5.4 displays the result for regressing hours supply without any control variables. Contrary to the results in participation, the estimated average treatment effect of hours supplied is negative and we can see the increase in hours supplied ranges from 2 to 3 hours as an effect of Child Tax Credit. This estimated average treatment effect is statistically insignificant for all of the specifications at even 1% level. These results suggest that the expansion of the CTC provision does not cause statistically significant increase in hours supplied in the pooled population. The statistical significance should not be confused with the practical significance, the increase in hours supplied is economically important: an increase of 2 hours is almost a 5% increase relative to the legal weekly hours of 40. In all scenarios, we see positive effect on hours supplied of the Child Tax Credit expansion of the recipient families. The standard errors for the specifications do not change across columns. The standard errors are not large when compared to the coefficients hence justifying the insignificance of the results.

Similar to the previous analysis in participation, we can the estimate the same set of equations for the the husband of the household to account for the gender differences. The following table presents the results:

Table 5.5: Estimated results for hours supplied using the male sub-sample

	(1)	(2)	(3)	(4)
VARIABLES	DID	DID: Regression Adjustment	DID: State FE	DID: State FE and education adjusted
treat_eff	2.112	2.132	1.882	1.828
	(3.300)	(3.253)	(3.250)	(3.247)
age		-5.909***	-5.952***	-5.022***
		(0.258)	(0.258)	(0.261)
race		0.00573	0.00267	0.00958***
		(0.00349)	(0.00359)	(0.00363)
nchild		-4.951***	-4.498***	-3.939***
		(0.617)	(0.619)	(0.623)
Constant	163.9***	533.4***	505.1***	465.1***
	(1.443)	(7.317)	(7.894)	(16.07)
Observations	348,120	348,120	348,120	348,120
R-squared	0.001	0.023	0.027	0.029
State FE	NO	NO	YES	YES
Regression Adjustment	No	YES	YES	YES
Education Control	No	No	No	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The results for the male sample is consistent with the estimates obtained using the pooled sample. The results suggest that apart from the small reduction in labor participation the expansion of CTC provision induces a positive change in hours supplied of male sub-sample. Over our the specifications, we see an increase in hours supplied about 2 hours. However, these results are statistically insignificant at usual levels. Again, we argue that the practical significance of these results are not to be overlooked. The U.S. Bureau of Labor Statistics estimated that the average weekly hours worked in 2020 is about 34 hours. Then our results indicate that there is about a 5% to 6% change in weekly hours supplied relative to the baseline hour of 34.

We can the estimate the same set of equations for female sub-sample and the following table presents the results:

Table 5.6: Estimated results for hours supplied using the female sub-sample

	(1)	(2)	(3)	(4)
VARIABLES	DID	DID: Regression Adjustment	DID: State FE	DID: State FE and education adjusted
treat_eff	3.580	3.076	2.915	2.523
	(3.406)	(3.352)	(3.349)	(3.345)
age		-6.503***	-6.560***	-5.437***
		(0.276)	(0.276)	(0.284)
race		0.0165***	0.0111***	0.0137***
		(0.00364)	(0.00378)	(0.00379)
nchild		-0.00838	0.365	-0.176
		(0.689)	(0.690)	(0.693)
Constant	146.6***	500.3***	478.7***	477.5***
	(1.459)	(7.410)	(8.023)	(22.59)
Observations	315,299	315,299	315,299	315,299
R-squared	0.001	0.022	0.026	0.027
State FE	NO	NO	YES	YES
Regression Adjustment	No	YES	YES	YES
Education Control	No	No	No	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Over our the specifications, we see an increase in hours supplied from 2 to 3 hours. However, these results are not statistically significant at any usual levels. The magnitude of this change in . The U.S. Bureau of Labor Statistics estimated that the average weekly hours worked in 2020 is about 34 hours. Then our results indicate that there is about a 5% to 6% change in weekly hours supplied relative to the baseline hour of 34.

5.2 Event Study(ES) Models: the dynamic treatment effect

A natural extension to the baseline Diff-in-Diff analysis is the class of ES models. There are several reasons why we wish to estimate an ES model in addition to the baseline Diff-in-Diff analysis. First, ES models can capture the dynamics of treatment effect before and after the event data that offers much more details. Together with the dynamic treatment effect, the ES models can also provide a convenient graphical illustration of the dynamic treatment effect. This characteristic can be greatly helpful in testing the identification assumptions of the baseline Diff-in-Diff model. That is, we can easily test and graphically demonstrate whether the parallel (pre)trend assumption holds. The flexible structure of ES models allows for the model to incorporate more realistic features such as variable intensity of treatment [28] and multiple treatments [56]. In these section, we will first formulate the set of ES models that we estimate and compare the results to the Diff-in-Diff models. Then we will discuss the graphical results and validate the identification assumption. Finally, we wish to incorporate the varying level of treatment intensity and address the heterogeneous treatment effect between male and female following the strategies suggested by [36], [58], and [19].

The standard ES Model can be estimated by the following equation:

$$y_{i,t} = \sum_{j \in \{-m,\ldots,n\}} \gamma_j D_{i,t-j} + \beta \boldsymbol{X}_{i,t} + \alpha_i + \delta_t + \epsilon_{i,t}. \tag{5.5}$$

Here, the dependent variable on the left hand side is $y_{i,t}$ can still be the weekly hours supplied and labor participation as before. Instead of the interaction of treatment status and post status in the baseline model, here we generated the event dummy variable $D_{i,t-j}$ for each observation. A distinction has been made between the event time j and the calendar time t by: $j = t - E_i$, where E_i is the calendar date where the event happens. Thus, the event time j can be interpreted as the periods since treatment. By design $j = 0$ at the event date. The event window $\{-m, \ldots, n\}$ is designated in ad hoc fashion and this choice depends both on the available data and theory, it will be discussed in more details. α_i and δ_t are standard two-way fixed effect controls and $\epsilon_{i,t}$ is the error term. We wish to control for fixed effects because we want these terms to capture confounds brought up by omitted variables. The event dummy variable $D_{i,t-j}$ is set to be 1 if at calendar time t the individual i experienced the event j periods ago. Thus, depending on the event time, the coefficient on the event dummies bares the interpretation of dynamic treatment effect. For all $j \geq 0$, γ_j captures the dynamics of the treatment effect, in particular, it shows how the treatment effect dissipates over time. γ_j such that $j < 0$ provides a good test for the identification assumption: if the identification assumption does not, then we should not to expect there to be any trends in these pre-event terms $\gamma_{j<0}$. The coefficient plot of γ_j over all event time provides a graphical illustration of the validity of experiment. Finally, similar to the previous Diff-in-Diff analysis, we wish to directly control for co-variates in the vector $\mathbf{X_{i,t}}$.

There are several ad hoc choices to make before estimating the event study model of this kind. The first choice is the selection of the event window $\{-m, \ldots, n\}$, in this case we will simply take the span of our data to be the event window. We will also following the convention in the literature over the choice of counter-factual reference period. That is, when estimating the event study model we will impose the restriction that $\gamma_{-1} = 0$ and all other $\gamma_{j \neq -1}$ are interpreted relative to the normalized reference period. Then the estimated coefficient of event dummy variables can be interpreted as compared to one period before the event, what is the impact. We will alter these choices and examine how sensitive the event study results are to these choices. In particular, another candidate for normalization is to choose a reference period $\{-k_1, \ldots, -k_2\}$ before the event date and set $\sum_{j\in\{-k_1,\ldots,-k_2\}} \gamma_j = 0$ and thus the estimated coefficients can be interpreted as the impact of the event relative to the average of reference period before the event. Having discussed the rudimentary regression restrictions, we can now estimate the model using our data with two candidates for normalization. Instead of presenting the table, we have the coefficient plot of the event dummies presented to summarize the results of the ES model and in the meanwhile to test our identification assumption.

With the popular normalization $\gamma_{-1} = 0$, we can see that the dynamics of the treatment effects of the expansion are not different from the baseline Diff-in-Diff result: for all $\gamma_j | j > 0$, we see that the estimated coefficients are not statistically different from at 5%. From the view of point estimates, we see that in the period of event, there was a positive response in labor participation and there were dips in two periods after the implementation of the CTC expansion and a gradual recovery in subsequent periods. However, since these point estimates are not statically significant, we cannot conclude that there were meaningful changes in labor force participation induced by the event.

We can thus focus on the pre-coefficients to test the validity of the quasi-experimental approach. The parallel pre-trend assumption requires the pre-treatment coefficients to be trend-less and any deviation from this pattern may suggest that the identification assumption is questionable. It is manifest that with the $\gamma_{-1} = 0$ normalization, the estimated coefficients have been largely trend-less, but there was a "dip" at $j = -3$ and $j = -2$. This result is indicative of the presence of the "Ashenfelter Dip" in labor participation. There are several plausible reasons that explain this dip: partly this is driven by the normalization choice and this can also be contributed to the presence of contemporaneous shocks that affect labor participation decisions prior to event.

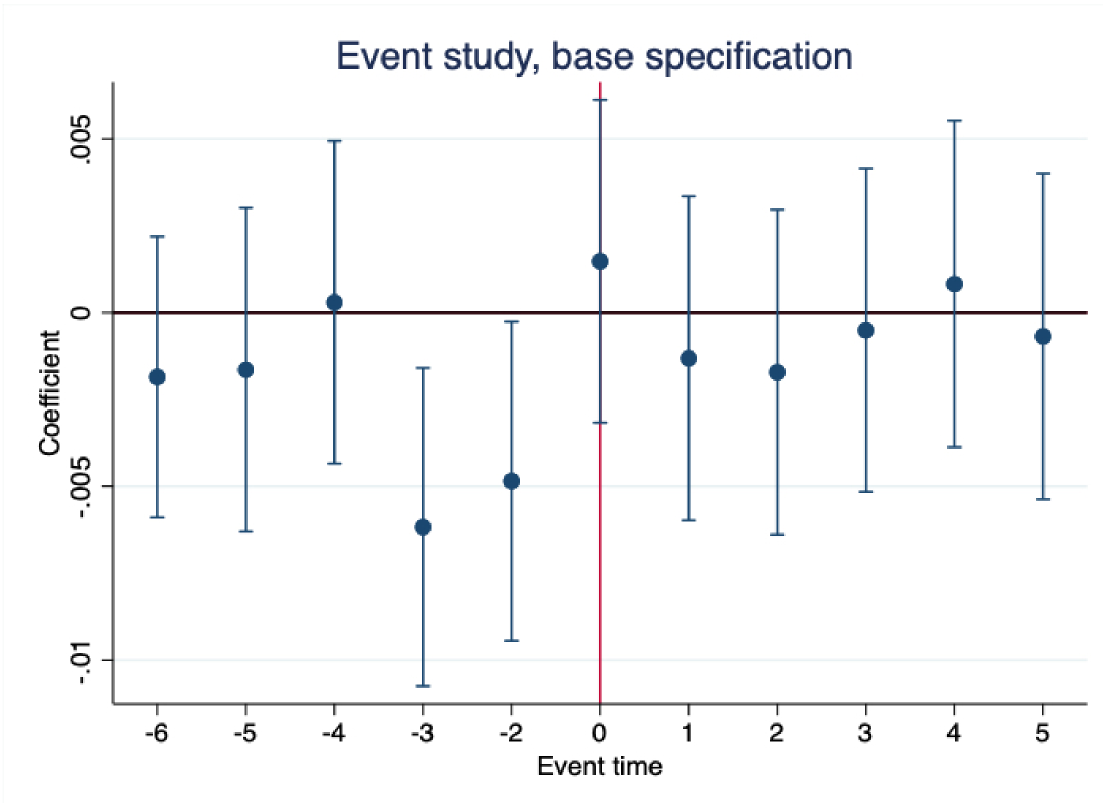

Figure 5.1: Coefficient plot for the γ_j of the base event study using the pooled sample and with normalization of $\gamma_{-1} = 0$

Instead of mechanically setting $\gamma_{-1} = 0$, we can choose a reference period before the event date and letting the average of coefficients in this reference period to be 0. We have chosen the reference period to be $\{-3, -2, -1\}$ and impose the restriction that $\gamma_{-1} + \gamma_{-2} + \gamma_{-3} = 0$. The following figure shows the coefficient plots for the initial and the alternative normalization specifications.

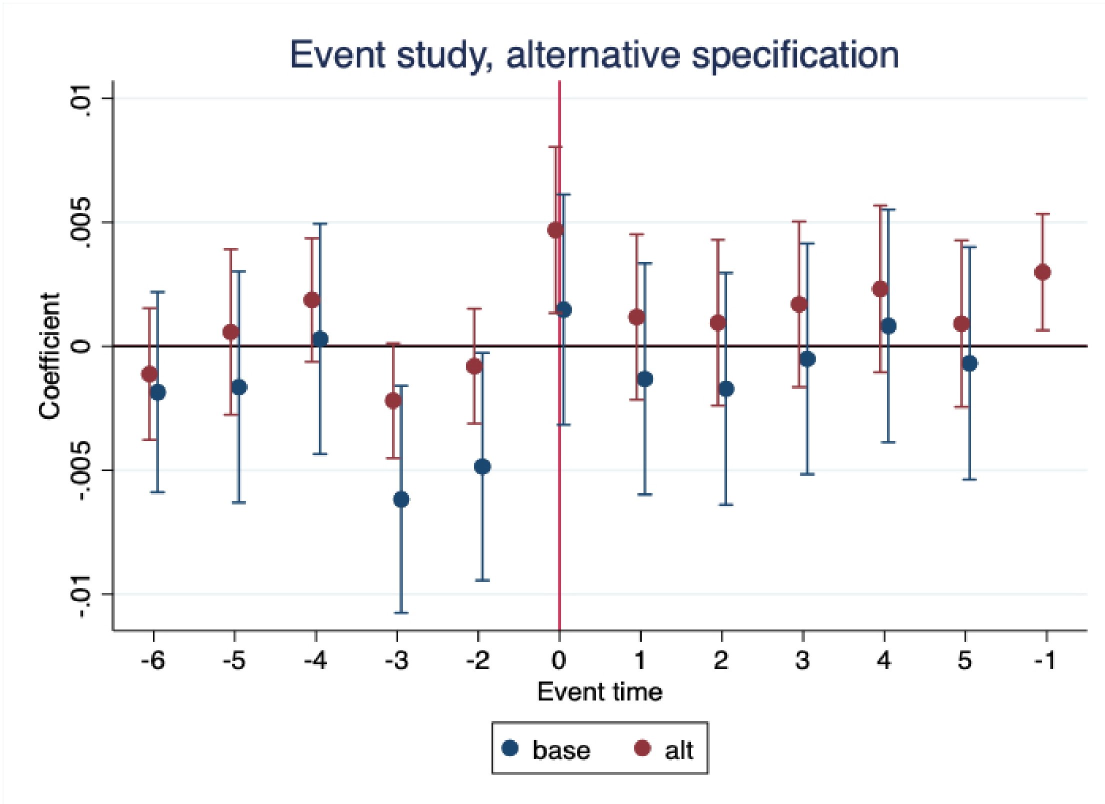

Figure 5.2: Coefficient plot for the γ_j of event study using the pooled sample and with normalization of $\gamma_{-1} + \gamma_{-2} + \gamma_{-3} = 0$ is presented in red and the coefficient plot using $\gamma_{-1} = 0$ is plotted in blue; the event coefficient for -1 is only visualized for the alternative normalization since it is mechanically set to 0 under the base normalization.

There are several observations from this figure. First, by imposing the alternative normalization assumption, the entire stream of event coefficients is shifted up since the counter-factual is different. For our case, the dynamic treatment effect is larger when the reference period is $j \in \{-3, -2, -1\}$ compared to setting merely $\gamma_{-1} = 0$. In addition, by extending the counter-factual pre-event reference period to more periods before the event, the confidence interval of the point estimates shrinks relative to the base normalization. Thus, we may conclude that regardless of the counter-factuals, the estimated impact of the event on labor participation is not statistically significant. This result is consistent with theoretical prediction of conditional cash transfer programs. Furthermore, the alternative normalization further corroborates the robustness of the results. The alternative normalization shifts the point estimates up and shrinks the confidence interval of the point estimates. The post event coefficients under this restriction is more likely to be significantly greater than 0 relative to the base specification. Our estimates, however, suggest that there are no significant increase in labor participation even with this alternative normalization.

Apart from the visualization of the dynamics of treatment effect, the final step is to quantitatively test the identification assumption that the pre-event coefficients are trend-less. This can be easily carried out by a T-test. We first perform the test to the base normalization. The p-value for the t-statistic is 0.79. Thus we fail to reject the null hypothesis that the pre-event treatment effect is trend-less. Similarly, we can do the same for the alternative normalization. With a p-value of 0.25, we again fail to reject the null hypothesis that the estimated pre-event coefficient is trend-less. Therefore, the results we have from the event study models further lend support to the conclusions we obtain from the Diff-in-Diff analysis and provide a richer dynamic effect of the impact of the event on labor participation.

5.3 Implications and Analysis of the Reduced-Form Results

Having obtained the initial estimates of the reduced-form models, it is possible to consider its policy implications. Over the span of 2021, we do not observe a significant response on both the extensive and the intensive margin of labor supply. From a policy-making perspective, the results suggest that the transfer from CTC kept and encouraged the parents participating in labor market through the credits provided to childcare. Depending on the specifications, it is possible to have a non-significant minor reduction in participation and hours supplied. A reduction in labor supplies may not be construed entirely negatively: if a parent decides to leave the labor force as the result of Child Tax Credit reform so that he/she may better take care of the children, the the overall long-run welfare consequence of such a policy may still have a positive impact on these household and the society as a whole despite such impact may not manifest in immediately obvious economic indicators.

Apart from the normative pondering on the welfare implications of the associated labor supply responses, it is also of our analytical interests to consider reasons why significant behavioral responses are absent from our data and analysis. There are two majors explanations to the observed responses: the first is the sticky responses and the second is heterogeneous responses.

5.3.1 Stick Labor Supply Response and Frictions

The sticky labor supply response explanation is based on the idea that due to various sources of frictions, individuals are unlikely and often times unable to make immediate adjustments in their labor supply decisions [32]. For this particular case, we suspect that the tax understanding frictions and the contractual frictions prevent the timely responses to be made. If the adjustments to the parameters of the Child Tax Credit (CTC) were implemented in July 2021, it is conceivable that households may not have been aware of these modifications until they file their 2021 taxes in April 2022. Consequently, any potential changes in labor supply resulting from the CTC may not be immediately apparent in the monthly data for 2021. The lagged responses have also been seen in several tax reforms in developed countries and we believe this is also one of the major reasons why immediate significant responses are not present in the 2021 monthly data [38]. The time lag between the implementation of the policy change and its reflection in the data can be attributed to the delay in households' awareness and reporting of the new CTC parameters.

The time lag between policy implementation and the observed responses can also be explained by contractual frictions. For instance, individuals bound by annual employment contracts may not have the ability to leave their job immediately following the month in which the CTC parameters changed. That is, the employment contract structures can separately provide a financial dis-incentive to individuals to adjust labor supply decisions. This concept is captured more generally by the presence of quasi-fixed costs in hiring. Quasi-fixed costs are often associated with hiring, training, and other investments in human capital that are not easily reversible or adjusted in the short term. They represent a form of adjustment frictions that firms face when making labor-related decisions. These costs can create barriers to labor market flexibility and affect firms' responsiveness to changes in demand or productivity [59].From a firm's perspective, it is costly to replace existing workers due to the presence of quasi-fixed costs and the firms will try to be less responsive and prevent employment changes on the extensive margin.

To address the delayed labor supply response and capture the dynamics of the policy impact, it is important to consider including multiple periods of data into 2022. This approach allows for a more comprehensive understanding of the long-term effects of policy interventions on labor supply decisions relative to some of the previous studies by Ananat and colleagues. By incorporating additional periods of data, we can capture the full dynamics of labor market adjustments, account for any time delays or frictions in response, and provide a more accurate assessment of the overall impact of the policy change.

By the particular nature of the policy and stylized facts in labor supply decisions, we argue that a more robust analysis can be achieved using more periods of data and a Quasi-Experimental analytical framework with more than one event per unit. That is, the initial policy enactment constitutes the first wave of treatment and tax filing on 15th April 2022 constitutes a second wave of treatment those who have incomplete knowledge of the CTC program. We shall discuss this in more details in the next section.

5.3.2 Heterogeneous Response and weighting

We discussed in the background section that the expected labor supply changes depend crucially on the segments on the budget constraints. This suggests that the treatment effect may depend on the observable socio-economic characteristics. That is, we suspect the policy to have a heterogeneous treatment effect on individuals. The estimated overall treatment effect from the quasi-experimental identification is a weighted average of the collection of heterogeneous treatment effects for each unit type [58]. One possible explanation for the insignificant overall treatment effect is that the weighting process diminishes some significant treatment effects for some unit-types. Two main streams of strategies are used when heterogeneous treatment effect is present: Sun and Abraham proposed an algorithm to find the weights; de Chaisemartin and D'Haultfoeuille suggested a two step procedure that first calculates the heterogeneous treatment effect and then recover the weighted overall treatment effect [18]. For the purpose of our study, we are only interested in identifying the heterogeneous treatment effect based on observable variables such as family income and education. We shall present the estimation results in the next section in more details.

CHAPTER 6

ROBUSTNESS CHECKS AND EXTENDED ANALYSIS

6.1 Non-Linear Response

Given that labor participation in labor statistics is coded as a binary variable, there are both economic and statistical reasons why we wish to estimate a non-linear model. Following the previous notations, we let y_{it} be the binary participation variable. To incorporate non-linearities, we can estimate the following models numerically using Maximum Likelihood Estimation(MLE):

$$P(y_{it} = 1) = \Phi(a + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 Treat_i Post_t + \theta \boldsymbol{X}_{ist}). \tag{6.1}$$

The following table presents the estimates:

Table 6.1: Non-Linear Labor force particimation

VARIABLES	(1) Probit: Female	(2) Logit: Female	(3) Probit: Male	(4) Logit: Male
treat_eff	0.0121	0.0232	0.00476	0.0171
	(0.0124)	(0.0211)	(0.0189)	(0.0368)
age	0.130***	0.230***	0.149***	0.265***
	(0.000842)	(0.00145)	(0.000880)	(0.00157)
race	-0.000139***	-0.000230***	-0.000209***	-0.000374***
	(1.39e-05)	(2.38e-05)	(1.54e-05)	(2.68e-05)
nchild	-0.0713***	-0.129***	0.138***	0.268***
	(0.00284)	(0.00492)	(0.00441)	(0.00836)
Constant	-3.319***	-5.815***	-3.420***	-6.105***
	(0.0534)	(0.0923)	(0.0540)	(0.0946)
Observations	579,332	579,332	537,075	537,075

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the binary labor force participation, and the right hand side co-variates are the same as the previous linear models. The vector of demographic characteristic control co-variates are included to account for observable differences between the treated group and the control group. The variable $Treat_i$ is included to control for the unobservables that distinguish the treated population and the untreated population. The variable $Post_t$ captures over-time changes in labor participation. Under the probit specification, the conditional error distribution is assumed to follow a standard normal distribution. We also estimate the same specification with a different error distribution assumption. The set of logit models is estimated with the assumption that the conditional error follows a logistic distribution with $\mu = 0$ and $\sigma = 1$. Similar to our previous case, the directions of demographic controls are all as expected and these are highly significant at 1%: older individuals are more likely to participate and there exists a significant racial differences. An interesting observation in this estimation is that within a household, the female are less likely to reduce participation with more children of schooling ages whereas the male in the household are more likely to increase labor force participation. Again we are interested in the coefficients of the interaction term. Unlike the linear probability model, the non-linear model has partial (treatment) effects be individual specific and depend on the entire vector of demographic controls. The magnitudes of partial (treatment)effect are complicated, but the signs of the parameters indicate the directions of the partial effect. To compute the treatment effects, we can follow the following steps. First we predict two predicted probabilities for each observation using all co-variates, but we compute the first probability setting the interaction term equal to 1 and we compute the second probability setting the interaction term equal to 0. Then for each observation, we calculate the difference between these two predicted probabilities. Finally, we can define the treatment effect of the treated to be the averaged differences over all eligible households. For the female sub-sample, we find that with probit model, the wife in the household with children had a 0.27 percentage point higher probability of labor participation as a result of the policy change. The estimated impact for the male of the household is much lower: the wife in the household with children had a 0.06 percentage point higher probability of labor participation as a result of the policy change. The results from logit models are consistent with the probit results. This suggests that estimated policy labor supply impact are robust. Thus, over all our estimated models, we do not see a negative impact on labor force participation decisions. The estimated coefficient on the interaction terms are all positive and insignificant at usual significance levels. The predicted probabilities are also relatively small in its magnitudes. These results are consistent with our linear probability models.

Particularly, this model enables us to examine how the treatment effect varies with demographic features. Two important demographic factors we are interested here are the level of education and age. Here we are only considering the labor force participation response of the household wife since we know from the non-linear models that the magnitude of response of the males is limited. c Furthermore, as far as the labor force participation response is concerned, there is a pronounced gap in education. Education plays an important role. Over the span of working age, we see that regarding the effect of expanded CTC provisions, the sub-group with less than high school education are more responsive then the sub-group with more than high school education.

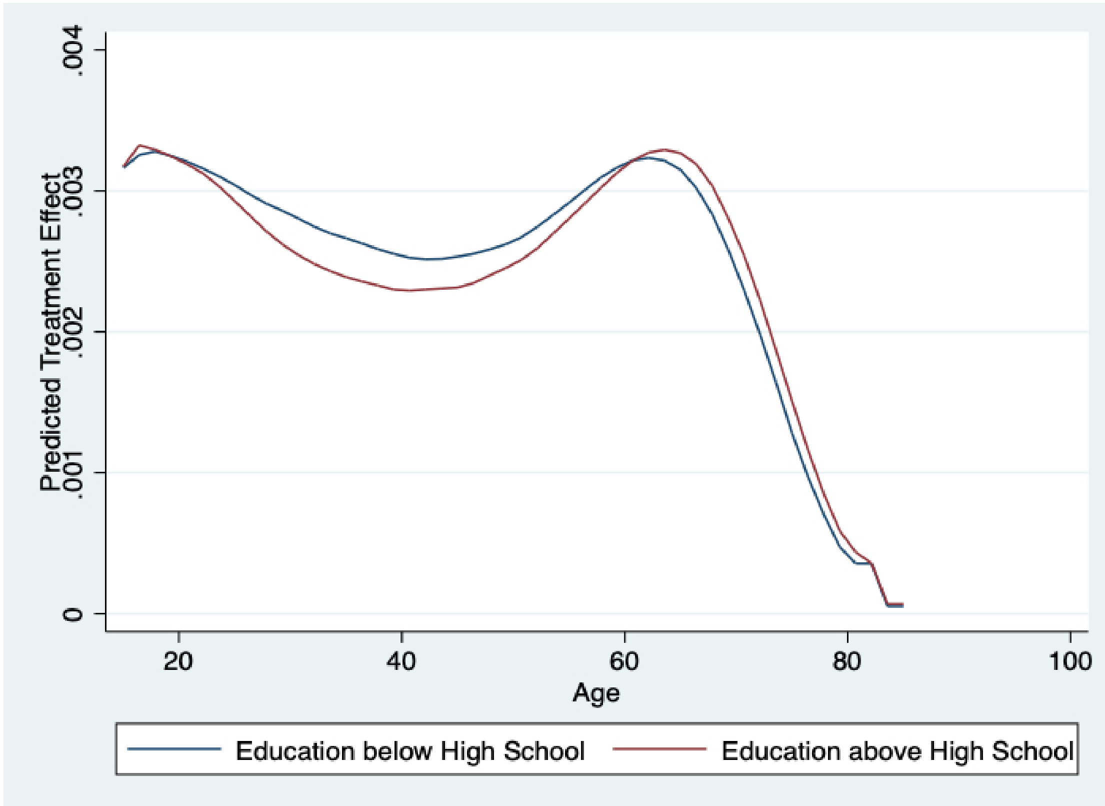

Figure 6.1: For the female sub-sample, the Predicted Treatment Effect varies as a non-linear function as demographic variables

Again, we can replicate this exercise to study how other demographic variables determine the response to CTC expansion. In this study, we consider the number of children of eligible age and the family income levels. Ruhm found that mothers with young children have lower labor force participation rates than those without children [55]. This can be attributed to the childcare cost considerations. Mothers with young children are likely to face a higher cost of childcare and thus this creates a quasi-fixed cost of participating in the labor force [43]. In principle, we should expect that families with more children are more responsive to the financial incentives offered by the changes in the CTC program.

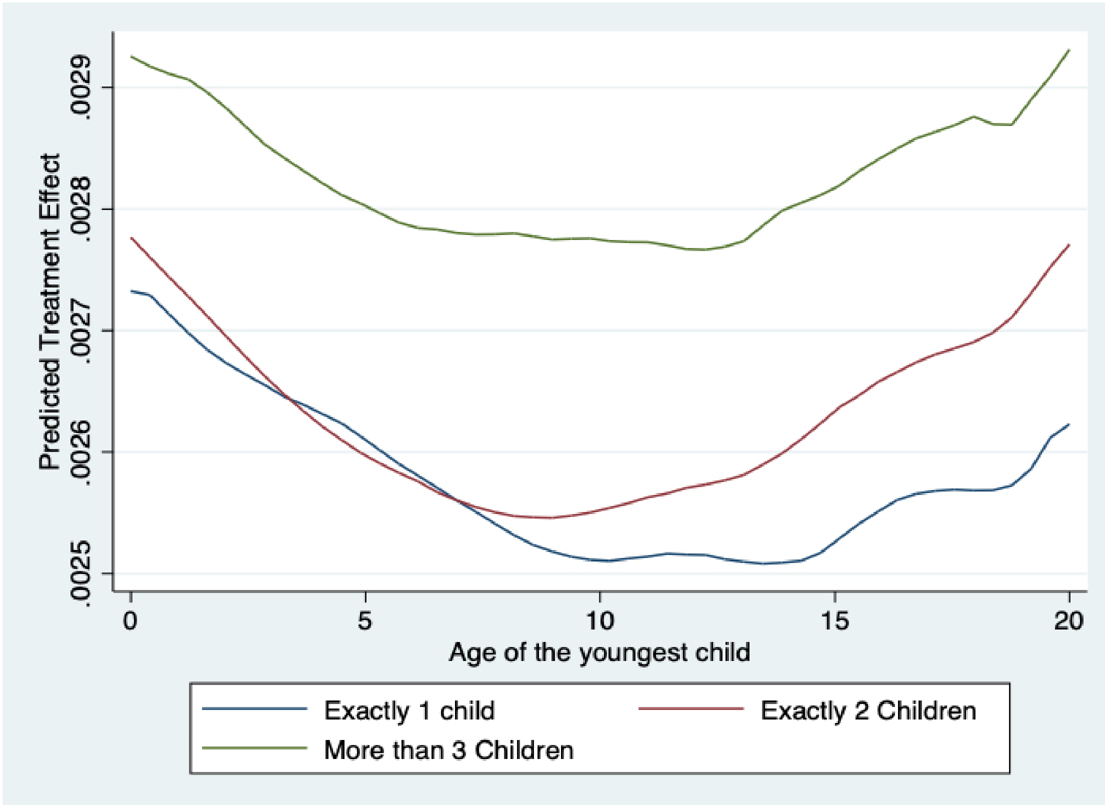

Figure 6.2: For the female sub-sample, we estimate the Predicted Treatment as a non-linear function of age and number of children.

Thus, the estimated treatment effect is consistent with our expectation. We observe that mothers with more children are more responsive to the policy changes in general. This is natural since the expanded size of transfer can help mothers to reduce childcare costs and thus the fixed costs of labor participation [51]. In addition, we assume that the childcare cost is a non-linear function of the the age of the children and thus indirectly affect the labor participation decisions of mothers. We will focus on the age of the youngest child in the household. The impact of children's age on participation probability follows an "U-Shaped" curve: we found that for all households(of different number of children), the labor participation responses are higher when the age of the youngest child are lower and when the youngest child approaches adulthood. Some plausible explanations apply: infants below 5 years old require more personalized care and attention whereas children approaching adulthood may require more structured and specialized activities and the level of activities may increase as they approach the age of 18 [29].

6.2 Delayed Response Analysis

To capture the dynamic effects of policy interventions on labor supply decisions and overcome some limitations of previous studies conducted by Ananat and colleagues, it is crucial to incorporate multiple periods of data into the analysis for the year 2022. This approach enables a more comprehensive understanding of the long-term consequences of the policy and provides a more accurate assessment of its overall impact. By including additional periods of data, we can capture the complete dynamics of labor market adjustments, consider any delays or obstacles in response, and enhance the robustness of the analysis.

Given the specific nature of the policy and the characteristic patterns observed in labor supply decisions, we argue that a more rigorous analysis can be achieved by utilizing an analytical framework that incorporates multiple events per unit, alongside an extended period of observation. This method has been used extensively used in the studies of labor market responses. Typical examples include multiple waves of layoff [40], and the continual hikes in EITC offering [21]. In this case, the initial implementation of the policy serves as the first wave of treatment, while the tax filing deadline on April 15, 2022, represents a second wave of treatment for those who have incomplete knowledge of the Child Tax Credit (CTC) program.

There are three main approaches to deal with plausible delayed responses: the first approach is to only consider the effect of the first event and suppress the impact of subsequent events; the second approach to incorporate multiple waves of treatments and delayed responses is to redefine the event dummy variables such that more than one event dummy variable can be non-zero at a time; the last approach is to duplicate the observation so that the data has observations by individual, event, and time [56]. For this section we will apply the first two approaches to analyze the labor supply responses because the first approach provide straightforward interpretations of the effect of the policy and the second approach is more robust to violations of the identification assumption. The first approach might introduce the selection issue as the never treated units' characteristics may change dynamically over time in unobserved ways and thus attach bias in the estimates of the first approach.

6.2.1 Analysis and Results

We estimate the first approach without considering the plausible subsequent event.

The coefficient plot of the event dummies is consistent with previous findings that there exists no immediate responses from the enactment of the policy in July 2021. With the extended monthly data, we can examine the overall dynamics of the treatment effect: from the policy period onwards a gradual recovery was seen despite the effects are not significant. The significant coefficients for event time 8 and 9 correspond to the treatment effect manifested in March and April in 2022.

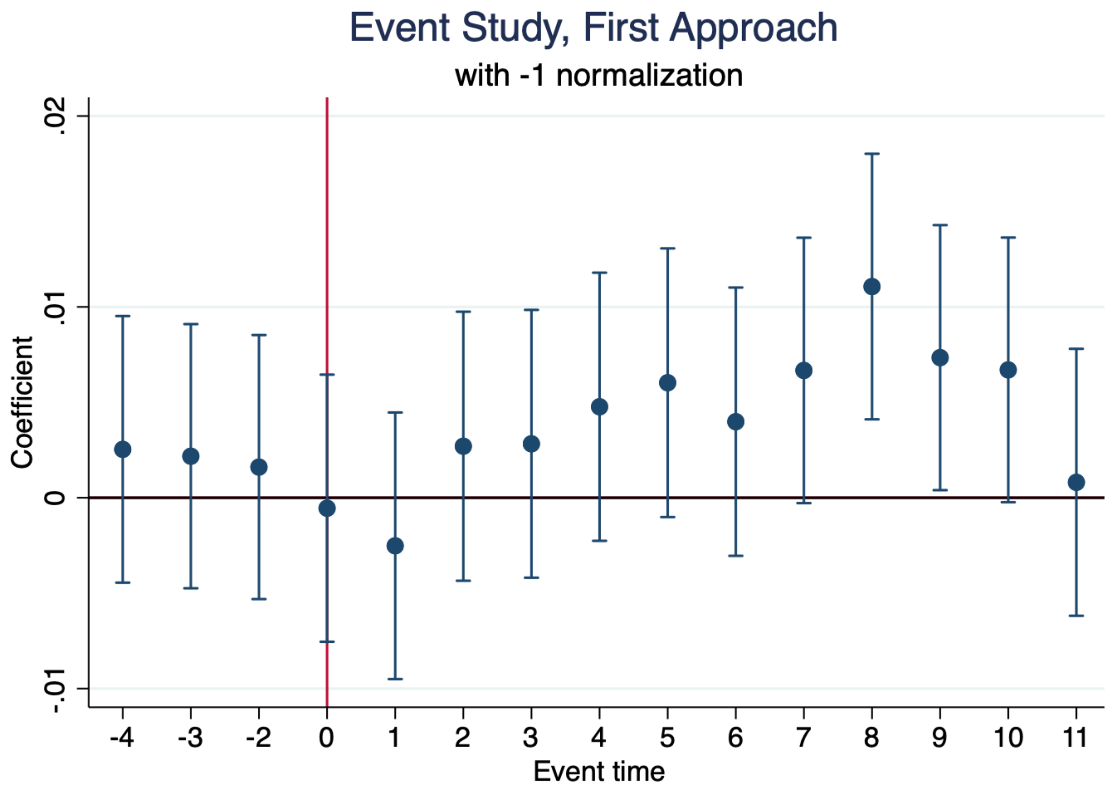

Figure 6.3: Delayed response using extended monthly data from 2021 to 2022

The direction is as expected: from the income and substitution effect analysis we would expect the participation to general increase following the program. Furthermore, the significant treatment effect in March and April 2022 confirms the hypothesis that there are frictions in the labor supply adjustment process. From the results, we claim further that the information/knowledge friction is the dominant explanation for the delayed labor supply response because the event dummy variables reverts to 0 beyond April. The positive response from March to April attests the presence of information friction: when the members of the household file the tax around March and April in the subsequent year they gain more information through the filing process and then adjust their labor supply decisions accordingly. If the contractual friction is an important cause for the delayed response, we would expect that the delayed treatment effect to remain significant beyond April as individuals are less constrained by the contract when they are further away from the event.

The first approach to delayed response confirms its existence, now we wish to utilize the second approach with multiple events since they are more robust to violations of the identification assumption and add the partial effect. To "turn on" multiple event dummy variable, this implies that $D_{i,t-j}$ is allowed to be one in whichever period when the event occurs for the same unit i. Thus we estimate the same set of equations as the base Event-Study case but with a set of re-defined event dummy variables and extended periods of data. Re-defining the event dummy variables changes the interpretation of the coefficients. The estimated coefficients now represent the partial effects of the event. Present event can have a potential impact on the likelihood of future events, thus the partial effect of present event should be interpreted as holding constant the impact of future events [39]. The partial effect has been given other names such as the Y-Channel-Only(YCO) effects indicating that spillovers through X-Covariate channels are not accounted [5].

The second model and the partial effects can thus be estimated.

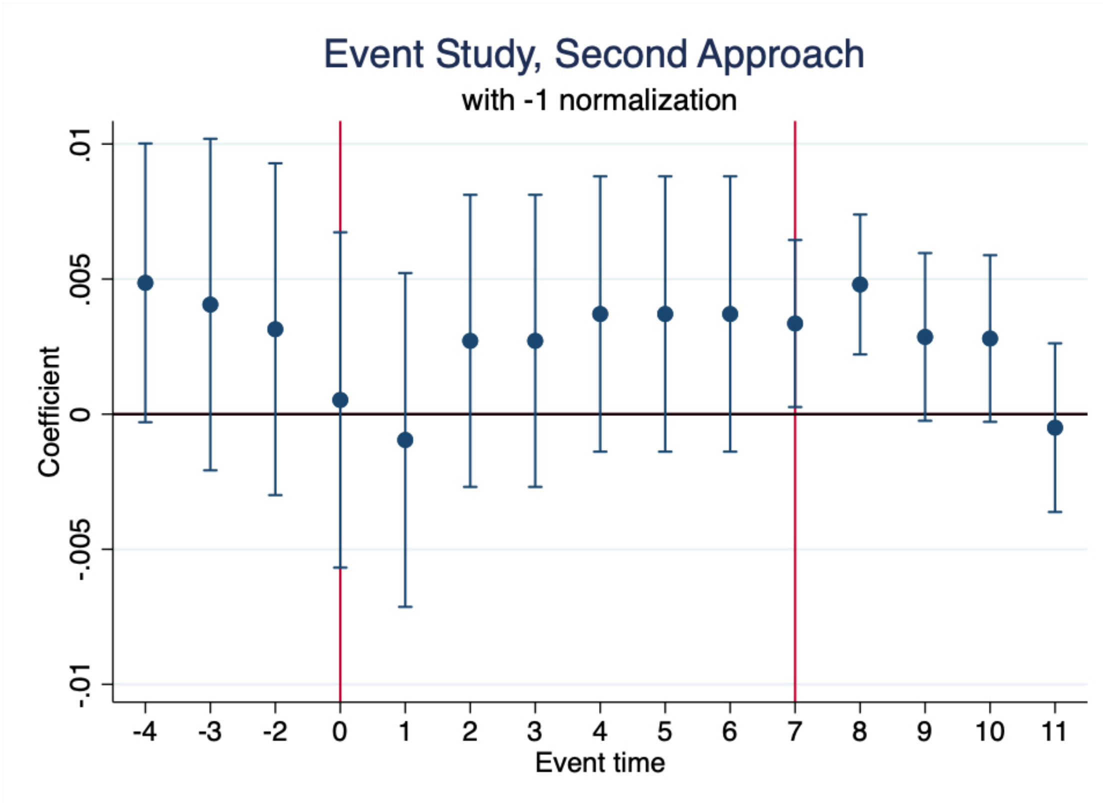

Figure 6.4: Partial Effects of the delayed response using extended monthly data from 2021 to 2022

In estimating the partial effects using the second approach with multiple events, we imposed additional restrictions. In particular, we pooled the coefficients after the first event so that we can obtain higher statistical power and narrower confidence intervals [47]. In this case we specified that the partial effect for $j = 2$ to be the same as $j = 3$ and the partial effect for $j = 4$ to be the same as $j = 5$ as suggested by Goodman-Bacon [27]. In addition, we modify the time of the second event. Instead of using the deadline of 15th April, we argue that it is more plausible that the second wave of treatment comes as earlier as February when individuals start to file their tax. The partial effect interpretation has an implicit assumption: the partial effect of the first and subsequent events are additively separable [56]. This assumption is unlikely to be violated for public programs such as tax hikes and transfer expansions, but if the dynamics of effects interact in non-linear ways, then the partial effect interpretation is no longer valid. By comparing the estimates of first approach, we can obtain an intuitive diagnostic for the additivity assumption.

The partial effects estimated by the second approach also support the delayed labor supply response caused by incomplete information about the program: the partial effects of the second wave of treatment appears to be positive and significant at 5% and reverts to 0 as time goes. This observation serves as further corroboration to our hypothesis of the delayed response since the dynamics of the estimated partial effects are consistent with that of the total effects estimated by the first approach. The separation of partial effects from the total effects offers a clearly defined impact of the subsequent information shock on the labor supply response. The analysis and conclusions drawn from both approaches are consistent with our hypothesis and produce statistically consistent dynamic causal estimates of the treatment effect of the policy reform. The observation that the total effect estimates are consistent with the more robust partial effect estimates adds another robustness check to the identification assumption of the quasi-experimental specification: the selection and the feature dynamics of the never treated control units are stable over time and the consistency of the our causal estimator can be further supported. The last piece of result is that the partial effects are additively separable indeed.

6.2.2 Policy implications

The analysis suggests that there exists a significant evidence for information frictions and these frictions set obstacles for timely labor supply adjustments. From a policy making perspective, it is important to address imperfect information. By addressing the issue of imperfect knowledge, policymakers can enhance the program's effectiveness and promote greater participation by reducing the exclusion errors. Obviously the ARP reform of CTC program did show effort in reducing the information friction by simplifying the program's structure and reducing administrative complexities to make it more comprehensible for the general public. It can be argued that further attempts to reduce the information friction might further promote the efficacy of the program. The information dissemination process can be improved by expanding the channels of information through various streams such as websites, helplines, and community outreach. It should be highlighted that the policy attempt to promote information access and knowledge should only be made after a thorough cost-benefit analysis. The implementation of various schemes to facilitate information circulation implies there are significant investment in installing these instruments. The policy maker should always balance the economic gains from diminished information frictions and the cost of reducing the degrees of imperfect information on the household side.

6.3 Do parents interact strategically?

The stark differences in the behavioral response left a number of questions to be asked. The first and perhaps the most important question we wish to ask is why is that the case. Therefore, a structural model is required to capture the response to incentives and the strategic interactions within the household. A husband and wife might engage in a strategic decision-making process when deciding their labor market participation, due to factors such as shared resources, household bargaining power, and complementarity or substitutability of their labor [62]. According to the economic theory of household decision-making, a family unit makes joint decisions regarding the allocation of resources, including labor market participation, based on their preferences and constraints [9]. This implies that the decision to work or not by one partner may be influenced by the other partner's decision, as it may affect the allocation of resources within the household. This strategic decision-making process can also be influenced by factors such as taxation policies and childcare costs, which can create incentives or disincentives for one partner to work or not. Therefore, a strategic decision-making approach can help to better understand the labor market participation decisions of couples and the corresponding responses to changes in relevant government policies.

The motivation to model this decision making process as a game is there exists externalities. We claim that childcare costs create an negative externaltiy for the labor participating decision in the household. If the cost of childcare is high, it may be more difficult for both player to work, as the cost of child care can offset the benefits of additional income [15]. Another source of externality is resources sharing. This is particularly relevant when the many of the variables in the CPS data is collected on household basis. For instance: the family income in the CPS data set represents the household total income and it is the same for both partners in the household.

The subject of our interests aligns with two areas of game theory: voluntary contribution to public goods and the common-pool resources. In many circumstances, the family income can be viewed as the shared resources within the household and the major contributor to this shared resources is the householders' labor wage. Therefore, the labor supply decision of each member can be viewed as a contribution decision to this shared resource [42]. However, it is not possible to model this purely as a public good model because this shared resource rivalrous and partially non-excludable. The shared resources depletes for one member if the other player overly consumes the available resources [50]. The shared resources are partially non-excludable, which means that it is difficult to exclude one member from accessing the resource. Despite various bargaining mechanisms that governs the access to the shared resources. strict exclusion is challenging, individuals may have open access to the resource. Such properties of the shared resources suggest that a common-pool resources game might be a more suitable modelling approach [30]. Compared with the traditional studies of common-pool resources that focus on the exploitation decisions, we are also interested in the contribution decisions via participation in formal labor market.

6.3.1 Data Construction

The data sample we constructed in the previous sections is of lesser utility for analysis because by the survey mechanism of the CPS, it is often times the case that there are more than two members in a household surveyed. Apart from including more periods of data, we wish to conduct our analysis with households composed of a (uniquely)married couple with or without children.

This allows us to eliminate the noises and biases in our estimates from other members in some big households where the married couples are not the only members.

To construct the date sample for the static game, we first trim the data in the previous section to obtain a sub-sample containing the decision and state variables of two unique partners in the household. We start with the female sub-sample from the previous section. First we keep the following variables: family income, age, sex, race, number of children, labor participation, education. Let the binary labor participation be the decision variable and others be relevant state variables. Each household is identified by an unique *serial* number. That is every individuals in a surveyed household share the same *serial* number. We only keep the observations without any repetitions in *serial* number. Finally, we denote the wife population as player j and rename the variables as var_j for all variables. Then we do the same for the husband population. We denote the husband population as player i and rename the variables as var_i for all variables. Then we can merge the wife data-set with the husband data-set to construct the sub-sample to estimate the static game. We use *serial* number as the unique identifier to merge two data-sets since the husband and wife share the same *serial* number and keep only exactly matched observations. This leaves us with 12,736 observations. In the language of strategic interactions, we as the econometrician observe the decision and state variables for $T = 12,736$ static games.

The next step in data preparation is to discretize variables in the sub-sample. The labor participation, age, sex, race, number of children is already discrete. State variables family income and education need more discretion since they are categorical. We will create a new discrete variable *highschool* that captures three groups: those with education attainments less than high school have *highschool* = 0, those with high school education have *highschool* = 1, and those with education attainments above high school *highschool* = 2. To discretize the family income, we need to consider relevant program parameters of the Child Tax Credit program. We can generate a discrete family income $faminc_d s$ such that it separate the households into three groups: families with $faminc_{ds} = 0$ have collective family income less than $150,000 and thus entitled to the maximum amount of CTC benefits; families with $faminc_{ds} = 1$ have collective family income above $150,000 and thus entitled to the partial or non CTC benefits. It should be noted that in the future steps of this study, it might be of our interest to have a more finely discretized family income levels and examine the impact on the structural estimates.

6.3.2 Identification and modelling

In this simple model, we assume the existence of a representative household consisting of a husband and a wife. Following the utility function defined by Shapiro and Stiglitz [57], we consider the payoff function to be a linear function of parameters drawn from the parameter space. The decision variables players make here is the binary labor participation decision. The state variables we include in the structural model include: age, sex, race, education, geographical information, family income, number of children. These variables are discrete or can be be discretized. We address the discretization procedures in the next section. These variables are recognized as public information: these information is observable to the other player.

We shall assume that each observation in the data set represents a static game being played by both partners. Due to the survey technique of the CPS data, it is difficult to extract a panel data set of large enough observations to conduct the structural modelling while ensuring the power of our statistical analysis. Therefore we chose to model this intra-household labor participation game as a static game. It should be noted that there are significant dynamic components and a dynamic game theoretical modelling is more suitable if relevant panel data can be gathered. We will apply and build on the structural econometric model of a static game that was developed by Bajari and colleagues to the subject of our interest [4].

We follow the standard notations stipulated by Bajari and colleagues: let a_i denote the action of player i and a_{-i} be the action of all the other players. We let $s_i \in S$ be the state variable. In vector notations, we let $\mathbf{a_t}$ be the vector containing the decisions of all players in the game and $\mathbf{s_t}$ be the vector containing the state variables of all players in the game t.

The first assumption we make is that the utility function is linear in parameters. The linearity in parameters assumptions implies that the utility functions can be expressed in forms of inner products. In particular, we let $\Phi(a_i, a_{-i}, s)$ to be the vector of basis functions.

Apart from the linearity in parameters assumption, we also assume that the utility function is additatively separable. This implies that then we can write the utility function U_i as a sum of deterministic component and a private information shock which is a stochastic disturbance. In particular we have:

$$U_i = U_i^0(a_i, a_{-i}, s_i; \theta) \;\; \textit{(Deterministic utility)} \tag{6.2}$$

$$+\epsilon_i(a_i) \;\; \textit{(private information shock)} \tag{6.3}$$

Another implicit assumption is that player i's deterministic utility is only affected by his own state s_i, not others. In this generic specification, θ is the parameter vector of the coefficients of arguments of the utility function we wish to estimate. These two assumptions combined allow us to use a compact notation for the deterministic utility as a linear combination of basis vectors $\Phi(a_i, a_{-i}, s)$: $U_i^0(a_i, a_{-i}, s; \theta) = \Phi(a_i, a_{-i}, s)'\theta$.

The term $\epsilon_i(a_i)$ is the private information shock. This can be viewed as yet another type of state variable that is only observable to player i and not observable to all other players and us the econometrician. In addition, we think of these shocks as dependent on the actions taken by player i and thus the action variable a_i enters as an argument. Since these state variables are only observable to the decision makers, we can model these as error terms in the utility function. In addition, we shall impose a distributional assumption such that each private information shock is independently and identically distributed with the extreme value distribution such that:

$$\epsilon_i(a_i) \overset{\text{iid}}{\sim} \tilde{f}(\epsilon_i(a_i)) \tag{6.4}$$

We can define the deterministic utility function. Suppose that player i choose leisure(not to participate) over work, then the utility function consists of the state variables and a_{-i}. To account for the asymmetries in payoff of this game, we should include individual specific state variables such as gender, education, and race so that we can identify the $\boldsymbol{\beta}$ parameter. Therefore, we can specify the deterministic utility function such that:

$$U_i^0(a_i, a_{-i}, s_i; \theta) = \beta_{1,i} nchild \cdot a_{-i} + \beta_{2,i} \cdot faminc \cdot a_{-i} + \theta' X_i, \quad if\ a_i = 0 \tag{6.5}$$

For the husband of the household, the deterministic utility function is a function of the number of children, the family income as well as the other partner's participation decision. Suppose that the other partner participates in labor market, we would expect that the childcare cost to fall more on the non-participating side and we expect that the parameter $\beta_{1,i}$ to be negative in signs. This interaction term captures the negative externality in labor participation decisions. Similarly, we would expect the parameter $\beta_{2,i}$ to be positive in signs since this interaction term captures the positive externality from the resources sharing schemes within this household. In addition, we expect the parameters to differ between players.

Suppose that player i choose work over leisure, then the deterministic utility function consists of the state variables and a_{-i} and a_i. We will follow the literature to allow for the interaction effect that captures the the impact of joint labor participation on the deterministic utility [3]. Then similar to the previous case with $a_i = 0$, we allow for the presence of cross effect through interacting decision variables with state variables [11]. Then we can specify the deterministic utility when $a_i = 1$:

$$U_i^0 = \beta_0 a_i a_{-i} + \beta_{2,i} \cdot faminc \cdot a_{-i} + \beta_3 nchild \cdot a_i + \theta' X_i, \quad if\ a_i = 1. \tag{6.6}$$

We assume that each player i has a policy function that outputs the discrete action conditional on the states and private information shock, we can then write:

$$a_i = \delta_i(\mathbf{s}, \epsilon_\mathbf{i}). \tag{6.7}$$

Note here that the other player's private information shock ϵ_{-i} does not enter as an argument in the policy function since this is not observable to player i when the binary decision is made. This policy function, however, is not known to either the other player or us as the econometrician. We assume the other player forms an expectation of the policy function and we call this the choice probability function or ex-ante policy function. This choice probability function is a function on observable state variables only and we denote it by σ_i. Given that we work with discrete data, we can use empirical averages as a non-parametric estimator such that:

$$\sigma_i(a_i = k|\mathbf{x} = \mathbf{m}) = \frac{\sum_\mathbf{t} \mathbb{1}(\mathbf{a_{it}} = \mathbf{k}\&\mathbf{x} = \mathbf{m})}{\sum_\mathbf{t} \mathbb{1}(\mathbf{x} = \mathbf{m})}. \tag{6.8}$$

The game id is indexed by t. We let k be the value of the decision variable and m be the value of the state variable. Then this non-parametric choice probability represents the conditional probability of player i choose $a_i = k$ when the observable state variables are $\mathbf{x} = \mathbf{m}$. The denominator is the total number of cases in which only the conditions $\mathbf{x} = \mathbf{m}$ is satisfied and the numerator is total number of cases when state variables satisfy $\mathbf{x} = \mathbf{m}$ and player i actually has decision variable $a_i = k$. Then we can take the product over each i and obtain $\sigma(\mathbf{a}|\mathbf{s}) = \prod_{\mathbf{i=1}}^{\mathbf{n}} \sigma_\mathbf{i}(\mathbf{a_i}|\mathbf{s})$. Since we know each player i does not know a_{-i} and ϵ_{-i}, the player i believes that other player's action is given by $\sigma_{-i}(a_{-i}|s) = \prod_{j\neq i} \sigma_j(\mathbf{a_j}|\mathbf{s})$.

Since each player i does not know a_{-i} and ϵ_{-i}, an expectation will be formed over these two terms based on the ex ante policy function. We let the expected utility be denoted by u_i^e. This can be written as :

$$u_i^e = u_{0i}^e + \epsilon_i(a_i) \tag{6.9}$$

In particular, the player i forms an expectation over the deterministic utility u_{0i}^e conditional on observable state variables and own action: $u_{0i}^e = \mathbb{E}\left(u_{0i}(a_i, a_{-i}, s_i, \theta | a_i, \mathbf{s})\right)$. In this expression, the only random term is a_{-i} and we need to take the expectation over a_{-i}. Since we have the expression for the ex ante policy function and discrete variables, this expectation is given by :

$$u_{0i}^e = \sum_{a_{-i}} u_{0i}(a_i, a_{-i}, s_i, \theta) \cdot \sigma_{-i}(a_{-i}|\mathbf{s}). \tag{6.10}$$

Suppose that we apply the linearity utility assumption and the compact notation of deterministic utility, we can transform the expected deterministic utility function:

$$u_{0i}^e = \Phi_i^e(a_i, s) = \mathbb{E}[\Phi_i(a_i, a_{-i}, s)|a_i, s] \tag{6.11}$$

$$= \sum_{a_{-i}} \Phi_i(a_i, a_{-i}, s) \prod_{j \neq i} \sigma_j(\mathbf{a_j}|\mathbf{s}) \tag{6.12}$$

A generic player i chooses his/her own decision variable a_i to maximize the expected utility u_i^e such that:

$$u_i^e(a_i, s, \epsilon_i; \theta) > u_i^e(a_i, s, \epsilon_i; \theta), \quad \forall j \neq i. \tag{6.13}$$

For finitely many state variables, the existence of the equilibrium can be shown using fixed point theorem [45]. Then the equilibrium optimal choice implies that the choice probability function has the following form:

$$\sigma_i(a_i|s) = Pr(\epsilon_i : u_i^e(a_i, s, \epsilon_i; \theta) > u_i^e(a_i, s, \epsilon_i; \theta), \quad \forall j \neq i). \tag{6.14}$$

Combining the equilibrium optimal choice probability, the linearity assumption, and the the iid extreme value distribution assumption of the private information shock, we can write the optimal choice probability in exponential functional forms:

$$\sigma_i(a_i|s, \Phi_i^e, \theta) = \frac{exp((\Phi(a_i, s)_i^e)'\theta)}{\sum_{k=0}^{K} exp((\Phi(k, s)_i^e)'\theta)}. \tag{6.15}$$

The discrete choice modelling and the extreme value distribution assumption made it easier for us to estimate the static game using a simple semi-parametric estimator [4].

The first step in the semi-parametric estimator for the structural parameters is to form a consistent estimator for the choice probabilities $\hat{\sigma}_i(a_i|s)$. There are several possible ways to do this. Bajari and colleagues suggested the sieves estimator [48], but other non-parametric estimators such as kernels and locally linear polynomials can also be used [25]. In our first stage estimation, we claim that the empirical average estimator presented in the previous section strikes a good balance between computational complexity and consistency. Having calculated the choice probability $\hat{\sigma}_i(a_i|s)$, we can then compute $\hat{\Phi}(a_i, s)$ by:

$$\hat{\Phi}(k, s)_i^e = \sum_{a_{-i}} \Phi_i(a_i = k, a_{-i}, s_i) \prod_{j \neq i} \hat{\sigma}(a_j|s). \tag{6.16}$$

Having obtained the consistent non-parametric empirical analogues, we can follow the equilibrium concept and compute the equilibrium choice probability evaluated at each possible structural parameter θ for each player.

The second step is to recover the structural parameter θ. This can be done in various ways: we first estimate the structural parameters using the Generalized Method of Moments(GMM) which is a more robust estimator; then we apply the Maximized Likelihood Estimator(MLE) to re-estimate the static game. We can test the robustness-variance trade-off by comparing the GMM and MLE estimators since the MLE estimator is more efficient when relevant assumptions holds.

To implement the second step, we first generate $y_{ikt} = 1$ if $a_{it} = k$ and 0 otherwise. Then we can collect this variable over players and actions to create a vector y_t. Similarly, we can collect the choice probability $\sigma_i(s_t, \hat{\Phi}, \theta)$ into a vector. The sample moment condition is defined to minimize the distance between the equilibrium choice probabilities and the observed decisions such that:

$$\frac{1}{T}\sum_t y_t - \sigma(s_t, \hat{\Phi}, \theta) = 0. \tag{6.17}$$

The structural parameters are obtained by minimizing the sample moment conditions numerically with identity matrix as the weighting matrix such that:

$$\hat{\theta}_{GMM} = \min_\theta \left(\frac{1}{T}\sum_t y_t - \sigma(s_t, \hat{\Phi}, \theta)\right)' I \left(\frac{1}{T}\sum_t y_t - \sigma(s_t, \hat{\Phi}, \theta)\right). \tag{6.18}$$

Let $B(s)$ be a matrix of state variables, $B(s)$ can be interacted with the original sample moment conditions to create additional moments to identify θ in cases of under-identification. Although the estimator obtained by the sample moment condition falls into the class of semi-parametric estimators studies by Newey and the formula for the asymptotic variances is provided [49], we will bootstrap for the standard errors for inferences. Bootstrapping for standard errors has desirable features in this context: it balances the computational efficiency and the distributional assumptions [20].

6.3.3 Results and Analysis

By minimizing the weighted moment conditions using the identity matrix as the weighting matrix, we can obtain the set of parameters that define the utility function of the husband and the wife of a household. The bootstrapped standard errors show that the estimated parameters are statistically significant at 1%.

$$\hat{\theta}_{GMM} = \begin{pmatrix} \beta_0 & = 0.4983 \\ \beta_{1,i} & = 0.5000 \\ \beta_{1,j} & = 0.4873 \\ \beta_{2,i} & = 0.5143 \\ \beta_{2,j} & = 0.4904 \\ \beta_3 & = 0.4822 \\ \beta_4 & = 0.5392 \\ \beta_5 & = 0.4335 \end{pmatrix} \qquad \hat{se}(\hat{\theta}) = \begin{pmatrix} 0.0025 \\ 0.0025 \\ 0.0025 \\ 0.0026 \\ 0.0025 \\ 0.0026 \\ 0.0024 \\ 0.0027 \end{pmatrix} \tag{6.19}$$

We can then recover the estimated utility function for each player of the intra-household game for each discrete choice of action. We can first recover the utility function of the husband(player i) when $a_i = lfp_i = 0$:

$$U_i^0 = 0.5nchild \cdot a_{-i} + 0.5143 \cdot faminc \cdot a_{-i} + 0.5392 * age_i + 0.4335 \cdot educ_{i,t}, \quad if\, a_i = 0 \tag{6.20}$$

For the husband of the household, the structural model confirms that there exists a significant strategic interaction within the household when deciding whether or not to participate in the formal labor market. The interaction term $faminc \cdot a_{-i}$ captures the positive externality from the resources sharing schemes within this household and the coefficient is indeed positive. As the other player(j) contribute to the shared resources by participating in the labor market, player i has incentive to free-ride on the shared resource. The coefficient on the individual specific controls are both positive and significant, this is consistent with our hypothesis. The coefficient on $nchild \cdot a_{-i}$ has mixed interpretations. On one hand this interaction term captures the childcare cost implied on the husband conditional on $a_i = 0$ and from a cost perspective we expect that as a_j increases from 0 to 1, the husband assumes the majority share of the childcare cost and thus coefficient $\beta_{2,i}$ is expected to be negative. On the other hand, however, one cannot ignore the intrinsic utility that parents derive from spending family times with their children. The latter effect is elegantly characterized as parental altruism by Nobel laureate in economics Gary Becker [6]. From the empirical estimation, it might be argued that in this model the parental altruism is the dominant impact in determining the utility function of the husband in a household.

Similar conclusion applies to the utility function of player j conditional on $a_j = 0$:

$$U_j^0 = 0.4873 nchild \cdot a_{-j} + 0.4904 \cdot faminc \cdot a_{-j} + 0.5392 * age_j + 0.4335 \cdot educ_j, \quad if\, a_j = 0. \tag{6.21}$$

The first observation from this result is that the coefficients for the individual specific state variables are the same as previous case since we assumed that the impact of these state variables is player-independent in this game. The second observation is that the direction/signs of interaction terms that capture the strategic interaction is the same as the previous case. This confirms that resources sharing creates a positive externality within the household and parental altruism in the parental utility function. A third observation is that the magnitude of the coefficients of the interaction terms are higher for player i(the husband). Since the dependent variable we study is discrete choice, the ordinal interpretation applies to the coefficients. That is, the magnitudes for both interaction terms are higher for the husband than the wife. This suggests that the benefits obtained from externality are higher for the husband than the wife in the household.

The structural parameters conditional on participating are also identified by the model:

$$U_i^0 = 0.4983 a_i a_{-i} + 0.4822 nchild \cdot a_i + 0.5143 \cdot faminc \cdot a_{-i} + \theta' X_i, \quad if\ a_i = 1. \quad (6.22)$$

Since we assumed the coefficients on the individual specific control is player-independent and independent of the action variable, these are the same as previous cases and are not presented here. The coefficient on the shared resources remains 0.5142 suggesting the presence of positive externality when the husband of the household participates in the formal labor force. The coefficient on $a_i \cdot a_j$ is 0.4983 suggests that there is a positive and significant interaction effect of the action variables of two players. This observation squares with positive externality we found in both $a_i = 1$ and $a_i = 0$. The coefficient of $nchild \cdot a_i$ is 0.4822 conditional on $a_i = 1$ further confirms our initial hypothesis that children provide positive contribution to parent's utility and this impact is positive in net.

The deterministic utility function player j(the wife) can be similarly captured:

$$U_j^0 = 0.4983 a_i a_j + 0.4822 nchild \cdot a_j + 0.4904 \cdot faminc \cdot a_i + \theta' X_j, \quad if\ a_j = 1. \quad (6.23)$$

By the modelling assumption, the utility function for the wife is almost identical to the husband's utility function apart from the coefficient of the interaction term capturing resources sharing. From our estimates of the structural parameter, it can be argued that there exits a significant element of strategic interaction when the wife and the husband of a household decide whether or not to participate in the formal labor market.

CHAPTER 7

POLICY IMPLICATION AND CONCLUSION

The welfare effects of Child Tax Expansion have been studied using simulated research with different conclusions: some studies finding a negative significant effect in labor force participation and some finding negligible negative effects [1]. Using real world data, this study uses difference in difference for the entire sample, and male and female sub-samples with additional robustness checks using demographic controls and doubly robust difference-in-differences model. Our findings suggest that there is no negative effect of CTC on labor supply, which is hypothesized in previous literature.

For the entire sample and women, the CTC expansion did not show any negative impact in the labor supply, which could be useful considering the implications of such an expansionary policy. There is always a threat of changing labor when policies regarding expansionary allotments are enacted with some literature agreeing that it positively affects women's labor force participation [24], [46] and negatively affect's men's labor force participation [24]. Our findings suggest we cannot fully agree on previous findings about labor supply when expansionary policies are made and thus provide a new perspective to the issue.

This study is based on initial analysis done regarding the Child Tax Credit's effects on labor supply. The findings lead to no certain conclusion, but further studies need to be done to validate the results and check the heterogeneous affect of the policy across various groups in the sample. We could break the samples and check individual affects by age groups and also for samples with different income distribution. The policy responses could be different by income distribution because initially, low-income households were not receiving maximum CTC benefits whereas they could be more likely to change labor supply. So, assuming the labor supply differs by income as well as age. Because of working capacity and number of children, the policy outcomes are likely different. Therefore, this study needs to be extended for solid conclusions. Another natural extension of the baseline analysis presented in this study is to address the plausible concern over contemporaneous policy shocks by employing an Event Study framework while explicitly controlling for the magnitudes of other policy shocks to the sample of our focus.

At a high level, at least we could suggest that there is no significant negative impact of the CTC expansion in the labor supply.

APPENDIX A

CHAPTER 1 OF APPENDIX

Appendix chapter 1 text goes here

BIBLIOGRAPHY

[1] Elizabeth Ananat, Benjamin Glasner, Christal Hamilton, and Zachary Parolin. Effects of the expanded child tax credit on employment outcomes: Evidence from real-world data from april to december 2021. Working Paper 29823, National Bureau of Economic Research, March 2022.

[2] Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2009.

[3] Patricia F Apps and Ray Rees. Collective labor supply and household production. *Journal of political Economy*, 105(1):178–190, 1997.

[4] Patrick Bajari, Han Hong, John Krainer, and Denis Nekipelov. Estimating static models of strategic interactions. *Journal of Business & Economic Statistics*, 28(4):469–482, 2010.

[5] Gaetano Basso, Douglas L Miller, and Jessamyn Schaller. Dynamic treatment effects for empirical microeconomists: Local projections and quasi-experimental research designs, 2022.

[6] Gary S Becker. *A treatise on the family: Enlarged edition*. Harvard university press, 1991.

[7] Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275, 2004.

[8] Antoine Bozio. Extensive and intensive margins of labour supply: working hours in the us, uk and france. *IZA Discussion Paper Series*, 6051, 2011.

[9] Martin Browning, Pierre-André Chiappori, and Yoram Weiss. *Economics of the Family*. Cambridge University Press, 2014.

[10] U.S. Census Bureau. Current population survey technical documentation. Technical report, U.S. Census Bureau, 2017.

[11] Elizabeth M Caucutt, Nezih Guner, and John Knowles. Why do women wait? matching, wage inequality, and the incentives for fertility delay. *Review of Economic Dynamics*, 5(4):815–855, 2002.

[12] CBPP. Policy basics: The child tax credit. Technical report, Center on Budget and Policy Priorities, 2021.

[13] Raj Chetty and Emmanuel Saez. Teaching the tax code: Earnings responses to an experiment with eitc recipients. *American Economic Journal: Applied Economics*, 5(1):1–31, 2013.

[14] Sophie Collyer, David Harris, and Christopher Wimer. Left behind: The one-third of children in families who earn too little to get the full child tax credit. *Poverty and Social Policy Brief*, 3(6), 2019.

[15] Rachel Connelly and Jean Kimmel. The effect of child care costs on the employment and welfare recipiency of single mothers. *Southern Economic Journal*, 69(3):498–519, 2003.

[16] Kevin Corinth, Bruce Meyer, Matthew Stadnicki, and Derek Wu. The anti-poverty, targeting, and labor supply effects of the proposed child tax credit expansion. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2021-115, 2021.

[17] Margot L Crandall-Hollick. *The Child Tax Credit: Proposed Expansion in the American Rescue Plan Act of 2021 (ARPA; HR 1319)*. Congressional Research Service, 2021.

[18] Clément De Chaisemartin and Xavier d'Haultfoeuille. Difference-in-differences estimators of intertemporal treatment effects. Technical report, National Bureau of Economic Research, 2022.

[19] Clément De Chaisemartin and Xavier d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996, 2020.

[20] Bradley Efron. *Bootstrap methods: another look at the jackknife*. Springer, 1992.

[21] Nada Eissa and Hilary Williamson Hoynes. The earned income tax credit and the labor supply of married couples. Working Paper 6856, National Bureau of Economic Research, December 1998.

[22] Nada Eissa and Hilary Williamson Hoynes. Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of public Economics*, 88(9-10):1931–1958, 2004.

[23] Nada Eissa, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner. Evaluation of four tax reforms in the united states: Labor supply and welfare effects for single mothers. *Journal of Public Economics*, 92(3-4):795–816, 2008.

[24] Nada Eissa and Jeffrey B Liebman. Labor supply response to the earned income tax credit. *The quarterly journal of economics*, 111(2):605–637, 1996.

[25] Jianqing Fan and Irene Gijbels. *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, volume 66. CRC Press, 1996.

[26] Naomi E Feldman, Peter Katuš čÁk, and Laura Kawano. Taxpayer confusion: Evidence from the child tax credit. *American Economic Review*, 106(3):807–835, 2016.

[27] Andrew Goodman-Bacon. Public insurance and mortality: evidence from medicaid implementation. *Journal of Political Economy*, 126(1):216–262, 2018.

[28] Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021.

[29] Elise Gould, Marcy Whitebook, Zane Mokhiber, and Lea JE Austin. Breaking the silence on early child care and education costs: A values-based budget for children, parents, and teachers in california. *Center for the Study of Child Care Employment*, 2019.

[30] Garrett Hardin. The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality. *science*, 162(3859):1243–1248, 1968.

[31] James J Heckman and Jeffrey A Smith. The pre-programme earnings dip and the determinants of participation in a social programme. implications for simple programme evaluation strategies. *The Economic Journal*, 109(457):313–348, 1999.

[32] Hilary Hoynes and Jesse Rothstein. Universal basic income in the united states and advanced countries. *Annual Review of Economics*, 11:929–958, 2019.

[33] Kosuke Imai, In Song Kim, and Erik H Wang. Matching methods for causal

inference with time-series cross-sectional data. *American Journal of Political Science*, 2021.

[34] Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

[35] IRS. Statistics for tax returns with the earned income tax credit (eitc). Technical report, Internal Revenue Service, 2020.

[36] Louis S Jacobson, Robert J LaLonde, and Daniel G Sullivan. Earnings losses of displaced workers. *The American economic review*, pages 685–709, 1993.

[37] Katherine M Keyes, Rebecca L Utz, Whitney Robinson, and Guohua Li. What is a cohort effect? comparison of three statistical methods for modeling cohort effects in obesity prevalence in the united states, 1971–2006. *Social science & medicine*, 70(7):1100–1108, 2010.

[38] Henrik Jacobsen Kleven, Camille Landais, Emmanuel Saez, and Esben Schultz. Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in denmark. *The Quarterly Journal of Economics*, 129(1):333–378, 2014.

[39] Pawel Krolikowski. Choosing a control group for displaced workers. *ILR Review*, 71(5):1232–1254, 2018.

[40] Jason M Lindo. Are children really inferior goods? evidence from displacement-driven income shocks. *Journal of Human Resources*, 45(2):301–327, 2010.

[41] Adam Looney and Monica Singhal. The effect of anticipated tax changes on intertemporal labor supply and the realization of taxable income, 2006.

[42] Shelly Lundberg and Robert A Pollak. Noncooperative bargaining models of marriage. *The American Economic Review*, 84(2):132–137, 1994.

[43] Shelly Lundberg and Elaina Rose. Parenthood and the earnings of married men and women. *Labour Economics*, 7(6):689–710, 2000.

[44] Chuck Marr, Chye-Ching Huang, Arloc Sherman, and Brandon Debot. Eitc and child tax credit promote work, reduce poverty, and support children's development, research finds. *Washington, DC: Center on Budget and Policy Priorities*, 2015.

[45] Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.

[46] Bruce D Meyer and Dan T Rosenbaum. Welfare, the earned income tax credit, and the labor supply of single mothers. *The quarterly journal of economics*, 116(3):1063–1114, 2001.

[47] Douglas L Miller. An introductory guide to event study models. *Journal of Economic Perspectives*, 37(2):203–230, 2023.

[48] Whitney K Newey. Semiparametric efficiency bounds. *Journal of applied econometrics*, 5(2):99–135, 1990.

[49] Whitney K Newey. The asymptotic variance of semiparametric estimators. *Econometrica: Journal of the Econometric Society*, pages 1349–1382, 1994.

[50] Elinor Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge university press, 1990.

[51] Stephen Roll, Leah Hamilton, and Yung Chun. Expanded child tax credit payments have not reduced employment. 2021.

[52] Jonathan Roth. Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–22, 2022.

[53] Donald B Rubin. Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American statistical association*, 75(371):591–593, 1980.

[54] Donald B Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

[55] Christopher J Ruhm. The economic consequences of parental leave mandates: Lessons from europe. *The quarterly journal of economics*, 113(1):285–317, 1998.

[56] Danielle H Sandler and Ryan Sandler. Multiple event studies in public finance and labor economics: A simulation study with applications. *Journal of Economic and Social Measurement*, 39(1-2):31–57, 2014.

[57] Carl Shapiro and Joseph E Stiglitz. Equilibrium unemployment as a worker discipline device. *The American economic review*, 74(3):433–444, 1984.

[58] Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.

[59] Robert H Topel and Michael P Ward. Job mobility and the careers of young men. *The Quarterly Journal of Economics*, 107(2):439–479, 1992.

[60] Burton Allen Weisbrod et al. The economics of poverty: an american paradox. *The economics of poverty: an American paradox.*, 1966.

[61] Philippe Wingender and Sara LaLumia. Income effects on maternal labor supply: Evidence from child-related tax benefits. *National Tax Journal*, 70(1):11–52, 2017.

[62] Matthias Wrede. The income splitting method: Is it good for both marriage partners? *German Economic Review*, 4(2):203–216, 2003.