**The Impact of Maternal Labor Supply on Subjective Well-Being: Correcting Endogeneity with a Bunching Design**

Shivani Gupta
University of Georgia
Department of Agricultural and Applied Economics
shivani.gupta@uga.edu

***Selected Paper prepared for presentation at the 2024 Agricultural & Applied Economics Association Annual Meeting, New Orleans, LA; July 28-30, 2024***

# The Impact of Maternal Labor Supply on Subjective Well-Being: Correcting Endogeneity with a Bunching Design

Shivani Gupta *

May 2024

**Abstract**

In recent decades, labor market developments with technological innovations and a desire for a better quality of life, especially post-COVID-19 pandemic have led individuals and policymakers to give more importance to the potential of flexible work. Despite a rise in female labor force participation, women still bear a disproportionate burden of childcare and domestic work, which comes at a cost of work-family conflict, resulting in poorer mental and physical health, and reduced well-being and life satisfaction. Using a nationally representative dataset for the Dutch population, this paper estimates the impact of maternal labor supply (weekly work hours) on their subjective well-being (mental health scores) with a special focus on heterogeneous effects based on the youngest child's age and the mother's labor intensity. We employ a novel identification strategy followed by Caetano, Caetano, and Nielsen (2023) using a control function approach with a bunching design for correcting the endogeneity. This approach leverages the potential bunching of observations for the maternal labor supply at zero hours of work, a pattern arising from the non-negativity constraint on time which helps isolate the effect of the unobservable confounders on the mental health outcomes to establish the causal treatment effects. Our results suggest that increasing the maternal labor supply negatively impacts their mental health. Part-time employment is associated with better maternal mental well-being with their mental health deteriorating if mothers who work part-time increase their work hours.

*University of Georgia, Dept. of Agricultural and Applied Economics. Corresponding Author: shivani.gupta@uga.edu

We don't find heterogeneity based on the youngest child's age indicating that part-time work could remain favorable for mothers' mental well-being, irrespective of their children's age. Our findings could inform policies promoting flexible work arrangements, which could help mothers balance their work and family responsibilities more effectively.

*Keywords:* Maternal labor supply, mental health, bunching, endogeneity, control function, part-time work

*JEL:* C10, D19, I10, J01

# 1 Introduction

Rapid technological advancement, labor market developments, and individual's desire for a better quality of life have led practitioners to assign greater importance to the potential of flexible work arrangements. The Netherlands known as the "first" (Visser, 2002) and the "only part-time economy in the world" (Freeman 1998) with a majority of women working part-time, has been acclaimed for its success in being a champion for balancing work and care, while also being ranked as one of the world's happiest countries (OECD, 2019).

In recent years, economists and social policymakers have increasingly recognized the importance of quantitatively analyzing subjective indicators (Layard, 2005) for effective policy evaluations (Veenhoven, 2002). Subjective well-being indicators such as mental health have been recognized to play an important role in the overall health of an individual as well as societal well-being (WHO, 2001). Over the past few decades, especially during the post-global economic crisis, researchers have progressively examined the relationship between economic factors and mental well-being (Strandh et al., 2013; Dagher, Chen, and Thomas, 2015) and empirical studies have now started to examine various economic factors affecting mental health, including poverty, inequality, household debt, and working hours (Powels, Siegers, and Vlasblom, 2008; Rupert et al. 2012; Piovani and Aydiner-Avsar, 2021; Ervin et al., 2023).

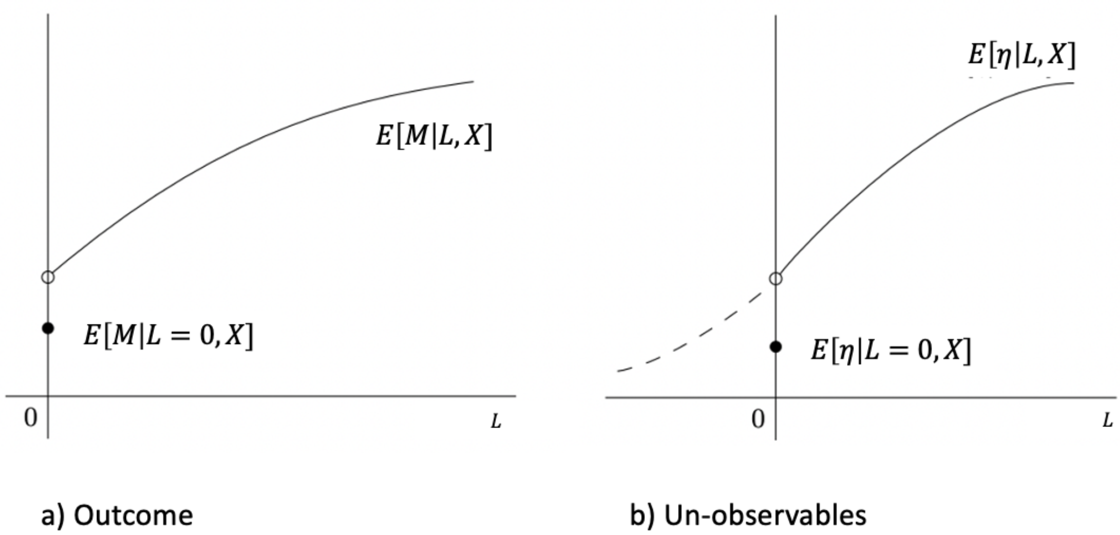
The choice of whether to work as well as how much to work, and its impact on health has been a widely debated issue. Moreover, most research studies exploring the effect of working time on subjective well-being are prone to selection bias and are unable to correct the problems of potential endogeneity which may hinder the identification of a true causal impact. The existing body of research on the effect of work time on subjective well-being shows mixed evidence. On one hand, working more can be a source of better mental health by bringing a sense of fulfillment, financial stability, and social connections (WHO, 2011), and on the other hand can be detrimental to mental well-being because of longer work hours, job insecurity, job stress, lack of on-the-job control and also, the double burden of productive (paid work) and reproductive (unpaid work like housework and child care) work especially, for women (Allen et al. 2000; Greenhaus et al., 2006; Minnotte and Yucel, 2018).

Over the last few decades, although, female labor force participation rates have increased in developed countries (Khoudja and Fleischmann, 2018), there still remains a huge gender gap in the global labor force participation rates among men and women (Campana, Gimenez-Nadal & Velilla, 2023). Many researchers suggest that these employment gaps could be due to gender norms and societal attributions about gender roles (Galván & Garcıa-Penalosa, 2018; Slotwinski and Roth, 2020). Women still face the added responsibilities of taking care of the child and doing domestic work (Wang and Coulter, 2019) while men's participation in unpaid household work still remains low (Sayer et al., 2009). Women, especially, mothers face a trade-off between time spent at work and time spent at home, and understanding this trade-off has become increasingly important as the maternal labor supply has increased (Fogli and Veldkamp, 2011; Khoudja and Fleischmann, 2018). Balancing work along with family responsibilities comes at a cost of work-family conflict, leading to poorer mental and physical health (Hammer et al., 2004; Greenhaus et al., 2006; Zhang et al., 2017; Inanc, 2018; Ervin et al., 2023), and reduced well-being and life satisfaction (Stoeva et al., 2002; Rupert et al., 2012). Considering these social obligations, working part-time can be a preference allowing mothers to have more family time (Balderson et al., 2021), which could help reduce their stressors and improve their mental health through work flexibilities (Oishi et al., 2015; Mache et al., 2020).

This paper estimates the effect of maternal labor supply on subjective well-being (mental health) outcomes using pooled panel data from 2008-2019 for mothers having young children from the Longitudinal Internet Studies for the Social Sciences (LISS), a representative survey of the Dutch population. The treatment variable, maternal labor supply, is calculated as the weekly hours worked, and the outcome variable is the Mental Health Inventory (MHI-5) score. Additionally, we control for maternal and child characteristics.

Estimating the effect of maternal labor supply on mental health can be challenging due to possible endogeneity and selection bias – reverse causality and correlation of maternal work hours with unobservables that directly affect mental health. Traditional research has tried to deal with these identification issues using methods like controlling for observables, individual fixed effects, and instrumental variables which require exclusion restrictions. We use a distinct approach to address the endogeneity and identify the causal impact of maternal labor supply on their mental health following a novel identification method by Caetano, Caetano, and Nielsen (2023) that involves bunching of the treatment variable which allows for building a control function. This method corrects for endogeneity by leveraging the potential bunching of observations for the maternal labor supply at zero hours, a pattern arising from the non-negativity constraint on time. The argument is that mothers working zero hours could have different unobservable confounders, but still have the same labor supply since they cannot choose negative hours to work. This helps to build a control function that isolates the causal impact of maternal labor supply by controlling the effect of the confounders on mental health. We, specifically, explore the heterogeneous effects based on the age of the youngest child and the mother's labor intensity to understand the decisions mothers make about how much to work.

Figure 1: Intuition about the control function approach




To explain how the control function approach works, consider the theoretical example shown in Figure 1. The left panel shows hypothetically how the outcome variable, mental health ($M$), varies with the treatment variable, hours worked ($L$). Controlling for the observed characteristics, $X$, the positive slope of the curve combines the treatment effect of working on maternal mental health and the endogeneity bias, the effect of mothers' unobservable characteristics $\eta$ affecting her choice to work, and her mental well-being. More specifically, it means that even holding $X$ constant, any variation in $M$ reflects both changes in the treatment $L$ and confounders $\eta$. As we can see, the left panel observes a discontinuity in the expected mental health at zero hours of work, i.e., the average mental health for mothers who do not work is discontinuously lower than for mothers who work a little more than zero hours. One explanation for this discontinuity could be that the treatment effect is discontinuous at zero. Now, we observe what happens in the right panel of Figure 1 which can give a more plausible explanation of this discontinuity in the mothers' expected mental health at zero hours of work. The figure shows how unobservable confounders that affect a mother's choice of how much to work varies with the actual work hours of the mother. This confounder could be a combination of factors such as mothers' skill, quality of time spent with the child, job insecurity, etc. that may affect the mother's choice to work and this could be specified as the "type" of the mother. The right panel shows that mothers who work slightly positive hours of work are of similar type on average, i.e., a mother who works 10 hours is similar to a mother who works 11 hours. The black dot on the figures shows the average value of the confounders for mothers who do not work at all, i.e., observation bunched at zero hours of work. While all these mothers have their hours of work lumped at zero with no variation in the treatment, in fact, these mothers are very different from each other in terms of their unobserved characteristics meaning variation in the confounders. Due to a non-negative constraint, these mothers cannot choose to work negative hours of work and have to choose a corner solution. This constraint is what plausibly explains the discontinuity observed in the outcome as seen in the left panel which implies that the discontinuity exists due to the unobservables alone without any contamination from changes in the treatment variable. Specifically, when we compare the outcomes at $L = 0$ and L just above zero, the difference in $L$ is negligible but the expected difference in $\eta$ is large which means the discontinuity exactly at $L = 0$ reflects changes in $\eta$ only without any changes in $L$. Hence, it is possible to identify the selection bias at the bunching point and indirectly, recover the treatment effect by controlling the effect of confounders on the outcome. However, this identification requires us to make assumptions about the size of the discontinuity on the confounder, since it is unobserved. This allows us to understand how far from indifference are mothers who want to work positive hours versus mothers who do not work at all. We talk about these assumptions in section 3.

Our findings reveal a negative impact of the increase in maternal labor supply on their mental health. We find heterogeneity present in our results depending on whether the mother participates in part-time work - the impact of work hours on maternal mental health becomes more negative for mothers who work part-time. Part-time employment is associated with better maternal mental well-being with their mental health deteriorating if mothers who work part-time increase their work hours. Finally, the presence of a young child aged 5 years or below doesn't significantly alter these relationships, indicating that part-time work remains advantageous for mothers, irrespective of their children's age. These findings could inform policies promoting flexible work arrangements, which could help women balance their work and family responsibilities more effectively. We also show that our main findings are robust to relaxed assumptions about the distribution of the unobservable confounders.

Our paper has two main contributions. The first contribution of this paper is to bring causal evidence of the impact of mothers' labor supply on their subjective well-being using a nationally representative dataset for the Dutch population. The second contribution is the identification strategy used in the study to address the endogeneity concerns and add to the nascent literature by estimating the average treatment effect using an entirely new source of variation leveraging the potential bunching of observations for the maternal labor supply.

The rest of this paper is organized as follows. Section 2 presents the data used for our analysis and Section 3 discusses our identification approach and specification assumptions. Further, sections 4 and 5 present the main impact results, robustness checks, and sensitivity analyses. Finally, we present our heterogeneity impacts in Section 6 and discuss our conclusions in the last section.

# 2 Data

We use data from the Longitudinal Internet Studies for the Social Sciences (LISS). LISS is a nationally representative longitudinal survey of 7500 individuals from 5,000 households of the Dutch population. The survey is conducted by CentER data, Tilburg using a probability-based recruitment method. LISS provides information on various measures of household/family characteristics and individual demographic and socio-economic characteristics, employment history, time allocation, physical health, and subjective well-being status.

For this study, we use the data from eleven waves – 2008 to 2019 – combining the three core modules: 'Family and Household', 'Work and Schooling', and 'Health' alongside the background data from all periods. Our final sample is a pooled cross-sectional data of 5866 mothers aged 21 to 55 years, who have a young child aged 18 years old or below living with them. Since there was no data collected on mental health in 2014, we dropped that year for the analysis.

Following the literature on subjective well-being, our study measures the main outcome variable using the Mental Health Inventory (MHI-5) score which is considered to be a reliable indicator of an individual's psychological distress (Thorsen et al. 2013). MHI-5 is a subset of the general health survey SF-36 and includes five questions referring to anxiety and depression. The survey records responses on a one-month recall scaled from 1 (never) to 6 (continuously) for questions including: "I felt very anxious", "I felt calm and peaceful", "I felt so down that nothing could cheer me up", "I felt depressed and gloomy" and "I felt happy". We created a standardized MHI-5 score from 0 (very bad) to 100 (very good) to measure the mental health of the mothers. Our treatment variable, maternal labor supply is measured as the total number of hours worked by the mother per week, calculated using an individual's actual working hours. To avoid any measurement error, we drop the data for respondents who are on maternity leave, mothers with any work disability, and mothers who are involuntarily unemployed (looking for a job). Finally, we also include a set of control variables based on the characteristics of the mother and the youngest child. These include mothers' age and age squared, and indicators for mothers' completed education: less than high school, vocational college and university degree, marital status, and bad habits: smoking (ever) and alcohol consumption (last 12 months). For the youngest child, we include indicators for their age: child's age 5 years or less, child's age 6-12 years, and child's age 13 years or more. We include year-fixed effects for our models.

Table 1 presents a summary statistics of our sample. Mothers have a mean mental health score (standardized) of 74 with a standard deviation of 15. For the treatment variable, the average number of hours worked per week currently is 18 hours with a standard deviation of 14. 22% of the mothers of our sample work zero hours per week. The rest of the summary describes the control variables used for the analysis. Mothers are 41 years old and work 20 hours a week, on average. More than half of the mothers are currently married (71%) and have completed some vocational college (63%) or university degree (9%). Among the sample of mothers, 81% consumed alcohol in the last 12 months, and 46% have smoked in their lifetime. Finally, in terms of children's age distribution, 32% of mothers have a youngest child aged 5 years or less, 37% of mothers have a youngest child aged 6-12 years, and 31% of mothers have a youngest child aged 13-18 years.

Table 1: Summary Statistics

| | Mean | Std. Dev. |
|---|---|---|
| **Outcome Variable** | | |
| Mental health score (Standardized MHI5) | 74.47 | 15.1 |
| **Treatment Variable** | | |
| Mothers' hours worked per week | 19.62 | 13.4 |
| **Bunching Variable** | | |
| Mother worked 0 hours per week | 0.22 | 0.41 |
| **Control Variables** | | |
| Mother's age | 41.41 | 6.79 |
| Mother's age squared | 1761.52 | 556.17 |
| Mother's age less than 30 years old | 0.04 | 0.19 |
| Mother's age 30 to 34 years old | 0.14 | 0.35 |
| Mother's age 35 to 39 years old | 0.21 | 0.41 |
| Mother's age 40 to 44 years old | 0.25 | 0.43 |
| Mother's age 45 years old or more | 0.36 | 0.48 |
| Mother's education less than high school | 0.26 | 0.44 |
| Mother's education vocational or some college | 0.63 | 0.48 |
| Mother's education university college | 0.09 | 0.28 |
| Mother is currently married | 0.72 | 0.45 |
| Mother has ever smoked | 0.46 | 0.49 |
| Mother consumed alcohol in last 12 months | 0.81 | 0.38 |
| Youngest child's age 5 years old or less | 0.32 | 0.46 |
| Youngest child's age 6 to 12 years old | 0.37 | 0.48 |
| Youngest child's age 13 to 18 years old | 0.31 | 0.46 |
| Year 2008 | 0.15 | 0.35 |
| Year 2009 | 0.10 | 0.31 |
| Year 2010 | 0.09 | 0.29 |
| Year 2011 | 0.08 | 0.27 |
| Year 2012 | 0.09 | 0.29 |
| Year 2013 | 0.09 | 0.28 |
| Year 2015 | 0.09 | 0.28 |
| Year 2016 | 0.08 | 0.26 |
| Year 2017 | 0.07 | 0.25 |
| Year 2018 | 0.08 | 0.26 |
| Year 2019 | 0.07 | 0.25 |
| *No. of observations* | 5866 | |

Source: LISS CentER data. Authors' calculations.

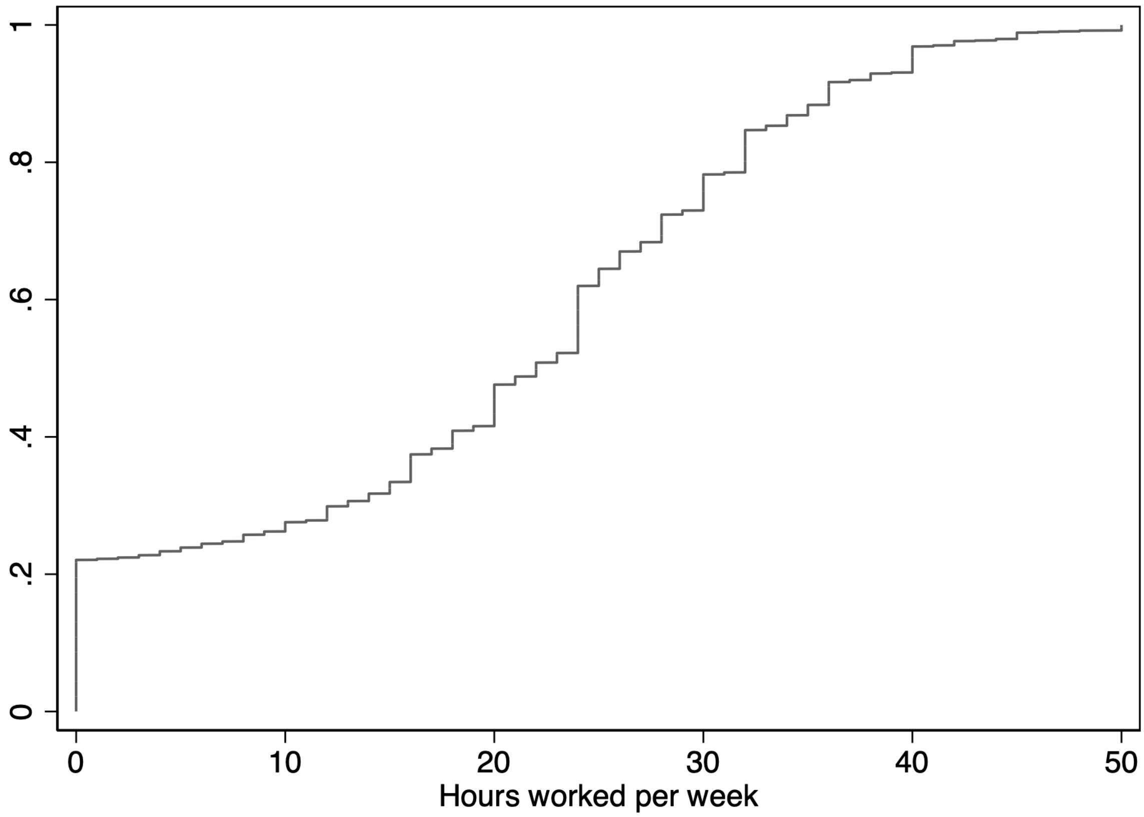
# 3 Empirical Strategy

In this section, we discuss our empirical approach for estimating the impact of maternal labor supply on their subjective well-being outcomes, i.e., mental health score (MHI-5). For this, we use a novel method of bunching estimation followed by Caetano, Caetano, and Nielsen (2023) to address the endogeneity using a control function approach. This leverages the potential bunching of observations for the maternal labor supply at zero hours, a pattern arising from the non-negativity constraint on time. The argument is that mothers working zero hours could have different unobservable confounders, but still have the same labor supply since they cannot choose to work negative hours. This helps us to build a control function that isolates the effects of confounders on mental health and hence, estimate the causal impact of maternal labor supply on mental health by controlling for the effect of the confounders.

## 3.1 Evidence of Bunching

In Figure 2, we show the cumulative distribution function (CDF) that illustrates the actual hours mothers spend on work ($L$) per week. It can be seen that there is a notable amount of bunching of observations, i.e., a concentration at $L = 0$ hours of work - about 22% of our sampled mothers do no work. This substantial bunching pattern persists across different types of mothers with children of different age distributions, i.e., mothers with the youngest child aged 5 years or less, mothers with the youngest child aged 6-12 years, and mothers with the youngest child aged 13-18 years as seen in Figure 2. This confirms the evidence of bunching in our sample and the evidence of bunching at $L = 0$ indicates that many mothers face a corner solution constraint when determining how many hours to work regardless of their type, i.e., how young their children are.

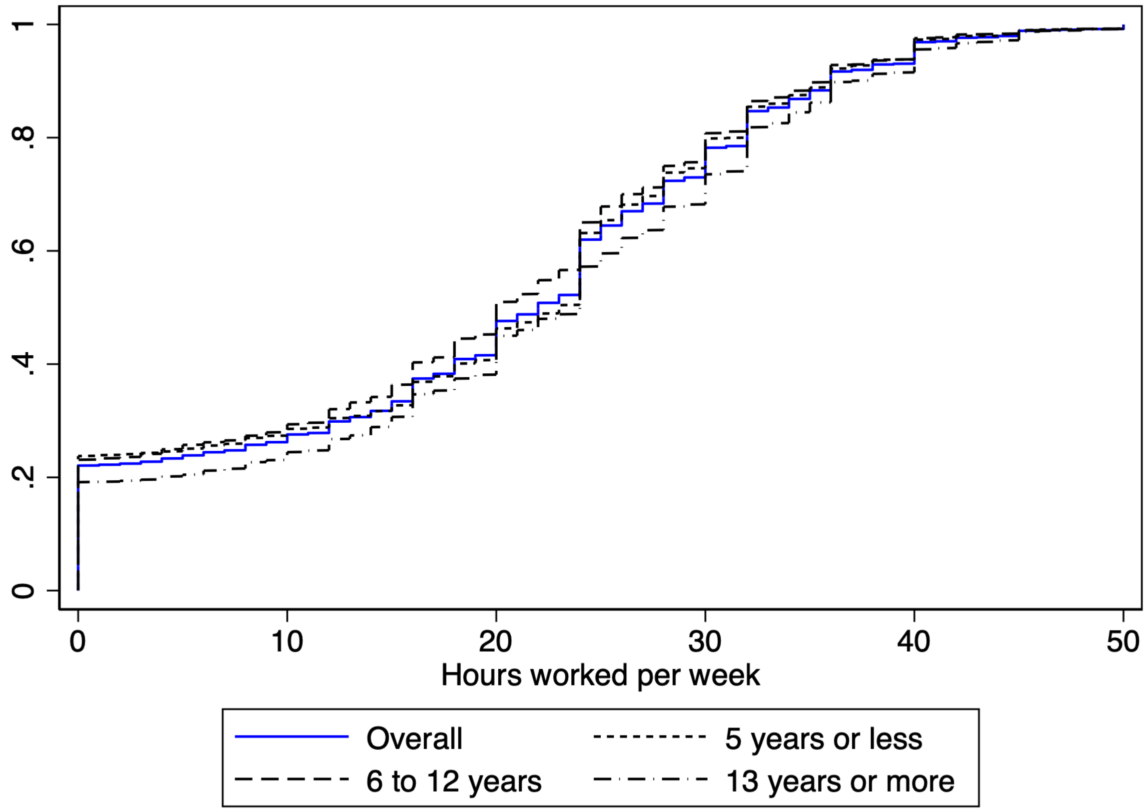
Figure 2: Evidence of Bunching: CDF for Hours worked per week




Note: This figure shows the estimated cumulative density function (CDF) of $0 \leq L \leq 50$ for the full sample.

Why does bunching exist at $L = 0$? To address this question effectively, it is essential to differentiate between the actual hours of work and the desired hours of work. Consider $L$ as the observed actual hours of work for the mother and $L^*$ represents the desired hours of work that the mother would choose if she had an unconstrained choice set (ability to choose negative hours if desired) which we cannot observe. Mothers deal with a restricted choice set while deciding how much to work because the actual hours $L$ they choose to work are limited to non-negative values so if the mother was really averse to working, she still may not choose negative work hours. While choosing how many hours to work, $L^*$ is influenced by a variety of observed and unobserved factors, which reflect the confounders influencing a mother's choice and can interpreted as an index of the "type" of the mother. Consequently, we define $X$ as the set of observed factors, while $\eta$ represents the rest of the

Figure 3: Evidence of Bunching: CDF for Hours worked per week by Child age distribution




Note: This figure shows the estimated cumulative density function (CDF) of $0 \leq L \leq 50$ for the full sample and samples with each age distribution of the youngest child.

unobservable confounding factors that affect mothers' choice of work hours. Therefore, we can write the desired hour of work as $L^* = h(X) + \eta$ such that $h(.)$ is non-parametric and $L = max\{0, L^*\},\ P(L^* < 0) > 0$.

Notice that mothers will choose their desired hour to work $L = L^*$ for $L^* \geq 0$ but will choose $L = 0$ for $L^* < 0$. This is because mothers who want to work a negative amount cannot do so. Therefore, we have two groups of mothers that have $L = 0$, first are the ones who are actually indifferent between working or not working, i.e., desire $L^* = 0$, and the other group is those mothers who are not indifferent, $L^* < 0$. Mothers who are at $L^* < 0$ are bound by a corner solution constraint $L = 0$ whereas mothers with $L^* = 0$ still have an interior solution with $L = 0$. We estimate the causal impact of maternal labor supply on mental health using this distinction between $L$ and $L^*$ at $L = 0$ when

$L^* < 0$ which helps identify the treatment effects.

To assess the impact of maternal labor supply on their mental health ($M$), we consider comparing $M$ among two groups of mothers with the same level of observed covariates $X$, when a mother works $L_0 > 0$ and when a mother increases her work hours from $L_0$ to $L_1 = L_0 + 1$ hours. This observed difference can be decomposed into average treatment effect and selection bias terms:

$$\underbrace{E\left[M \mid L = L_1, L^* = L_1\right] - E\left[M \mid L = L_0, L^* = L_0\right]}_{\text{What is Observed}} =$$
$$\underbrace{E\left[M \mid L = L_1, L^* = L_1\right] - E\left[M \mid L = L_0, L^* = L_1\right]}_{\text{Average Treatment Tffect}} +$$
$$\underbrace{E\left[M \mid L = L_0, L^* = L_1\right] - E\left[M \mid L = L_0, L^* = L_0\right]}_{\text{Selection Bias}}$$

Here, the observed difference in the average maternal mental health ($M$) compares mothers who choose $L = L_1$ to those with mothers choosing $L = L_0$, however, these two "types" of mothers differ in their unobservable factors: mothers with desired hours of $L^* = L_1$ and mothers with desired work $L^* = L_0$, respectively. By adding and subtracting the term $E[M|L = L_0, L^* = L_1, X]$, we are able to decompose the difference into the treatment effect and the selection bias term. The average treatment effect represents the causal impact of increasing work hours from $L_0$ to $L_1$ on maternal mental health, comparing the same "types" of mothers with desired work hours $L^* = L_1$ with identical observable attributes, $X$. Finally, the selection bias term accounts for the bias introduced due to the differences solely attributed to the unobservable factors $\eta$ between the different "types" of mothers: mothers working ($L^* = L_1$) and ($L^* = L_0$) hours, despite having the same observed work hours $L_0$. This factor represents the remaining source of variation in mothers' desired work hours of work beyond the observable factors.

To identify the causal effect of maternal labor supply $L$ on their mental health ($M$), traditional approaches shut off the selection bias term to identify treatment effects.

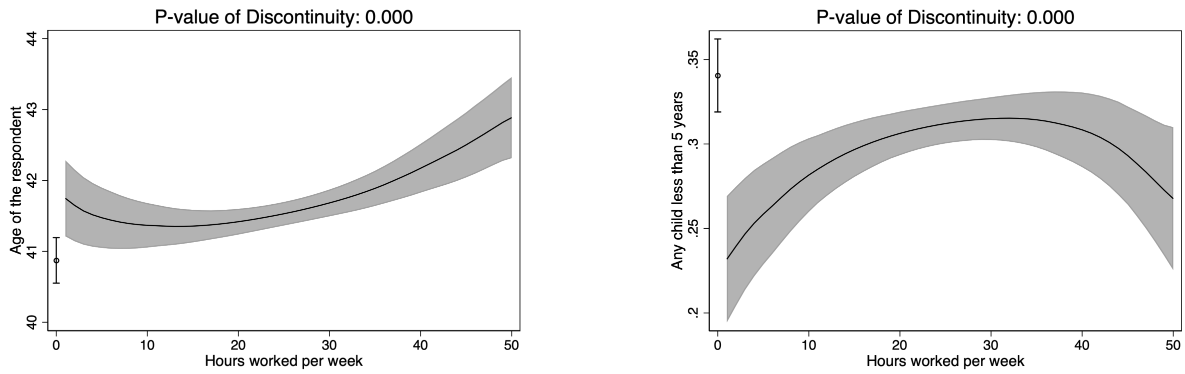
In this study, our empirical estimation uses a unique control function ("selection-on-unobservables") approach to indirectly estimate the treatment effect. This is done using a source of variation that identifies the selection bias term and by shutting down the treatment effect. The treatment effect can then be recovered by calculating the difference between the observed effect and the selection bias. In our analysis, we leverage the discontinuity in the outcome, maternal mental health ($M$) as we approach $L = 0$ from the right side, which we know is solely due to the unobservable factors $\eta$ for $L^*$. This approach enables us to identify the "selection bias" by isolating the variations in $L^*$ giving us the effect on $M$ attributed to different "types" of mothers through the effect of unobservables on $M$. The fundamental premise of this approach lies in the understanding that both $L$ and $L^*$ affect $M$, but $L$ changes discontinuously with $L^*$. To illustrate this concept, consider comparing mothers with $L = 0$ (indicating no work hours) and $L = l$ (slightly small positive value of $l > 0$). The observed differences in $M$ cannot be attributed to the average treatment effect, given the almost identical work hour choices ($L$) of these mothers. However, the types of these mothers may significantly differ. Therefore, the differences in mental health can only arise from variations in $L^*$, representing the different "types" of mothers. Mothers with $L = l > 0$ correspond to the type $L^* = l$, but those with $L = 0$ might not necessarily align with $L^* = 0$. While some mothers might be exactly indifferent between working some hours and not working at all ( $L^* = 0$), other mothers might be at a corner solution: their aversion to working is so strong that they would prefer negative hours, indicating that even though $L = 0$, $L^* < 0$. This results in a strictly negative average type among mothers at $L = 0$. Hence, at $L = 0$, there exists a significant variation among the types of mothers making such choices, and any discontinuity in average mental health at this point can only be attributed to discontinuities in these average types of mothers, i.e., the unobservable confounders. We present evidence indicating that some mothers at $L = 0$ must have $L^* < 0$, confirming the discontinuity in the average type at $L = 0$.

## 3.2 Evidence of Selection

Now, we show evidence that there is a selection-on-observables and selection-on-unobservables in our sample indicating that mothers who do not work are discontinuously different from mothers who work positive hours.

Figure 4 shows a local linear regression of key control variables on $L$ for the sample of mothers who work any positive amount of hours ($L > 0$) and additionally, it plots the average value of these variables for mothers who do not work at all ($L = 0$). The result in each panel shows that mothers who do not work are discontinuously different from mothers who work positive amounts of hours in all these observable characteristics. Specifically, in comparison with mothers who work positive hours, mothers who do not work at all are distinctly younger and are more likely to have a young child aged 5 years or less. These discontinuities are observed for several other observable characteristics, indicating that there exists selection-on-observables at $L = 0$. These discontinuities show that mothers who do not work tend to be selected with respect to covariates correlated to the mother's mental health.

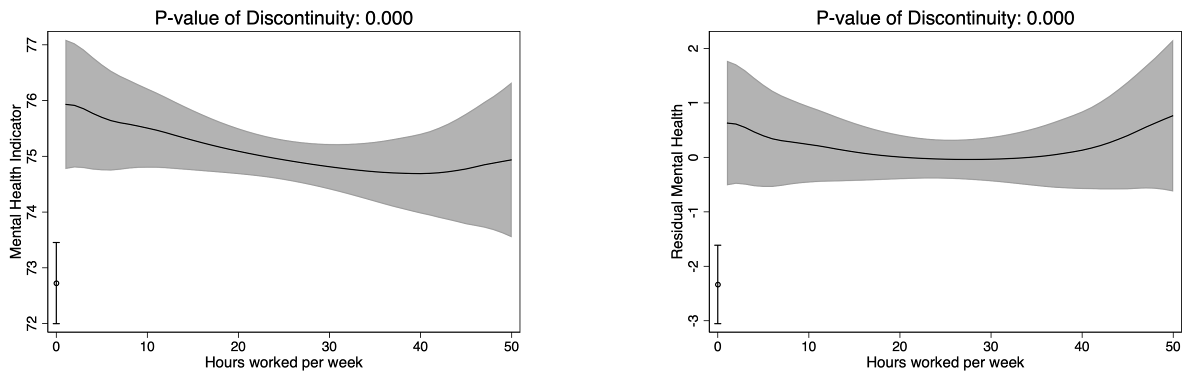
Figure 4: Evidence of Selection on Observables at $L = 0$




Note: (1) Each panel shows a plot of the local linear regression of two key observed covariates on mothers' actual hours worked per week $L$ along with the 95% confidence interval (2) The bandwidth is 15 hours. (3) At $L = 0$, the average along with the 95% confidence interval is also shown (4) The p-value of a test for whether there is a discontinuity at zero is shown in the header of each panel. Source: Authors' Calculations

Figure 5 shows the direct evidence of selection-on-unobservables using the local linear regression fit of the outcome variable (maternal mental health), $M$ on $L$ for the sample of mothers who work positive hours $L > 0$ and the average value of $M$ plotted for mothers who do not work at all $L = 0$. We can see that there exists a significant positive discontinuity suggesting evidence of positive selection with mothers who do not work tending to have lower mental health scores than mothers who work some positive hours. The right panel of the figure illustrates the "residualized" local linear fit based on the discontinuity test in Caetano, 2015, which shows that even after controlling for all observable characteristics, $X$, maternal mental health scores remain discontinuously lesser for mothers who do not work at all as compared to mothers who work positive hours. This indicates that the discontinuities in $M$ at $L = 0$ exist only due to unobservable confounders $\eta$ after controlling for observed covariates $X$ and hence, motivates our selection-on-unobservables approach in our paper to estimate the impact of $L$ on $M$ using a control function method. We explain this approach in our next section.

Figure 5: Evidence of Selection on Un-observables at $L = 0$




Note: (1) Analogous to Figure 4, however, instead of the covariates vertical axis plots the local linear regression of the outcome $M$ (maternal mental health) in the left panel and the "residualized" $M$ in the right panel on mothers' actual hours worked per week $L$ along with the 95% confidence interval. (2) The expected value of the variable at $L = 0$ is shown, along with its 95% confidence interval. (3) The p-value of a test for whether there is a discontinuity at zero is shown in the header of each panel. (4) The "residualized" variable uses covariates $X$ (described in Section 3) that enter non-parametrically in this regression, as described by Caetano (2015). In the regressions, $K = 20$ and $K_\delta = 1$ as in the main results shown in Section 4. Source: Authors' Calculations

### 3.3 Control Function Approach

We begin by defining the maternal mental health (M) as:

$$M = f(L, X; \beta) + \epsilon \tag{1}$$

For a homogeneous model,

$$M = \beta L + g(X) + \epsilon \tag{2}$$

Here, g(X) represents the non-parametric component, $L$ represents the number of hours worked by the mother, $X$ denotes a vector of pre-determined controls, and $\epsilon$ represents the unobserved error term. To allow for heterogeneity effect, our parametric function $f(\cdot; \beta)$ can be specified in different ways to understand the diverse effects of maternal labor supply $L$ on their mental health $M$. Our main goal remains the identification of $\beta$, regardless of the specific form. However, there is selection-on-unobservables selection after controlling for observable characteristics, arising from the endogeneity in $L$, i.e., $E[\epsilon|L, X] \neq 0$. Consequently, a regression of $M$ solely on $L$ and $X$ would result in a biased estimate of $\beta$.

As discussed previously, we know that mothers face a constrained choice to choose the amount of work they do because they cannot work negative hours. Hence, we can write the desired hour of work $L^*$ as a combination of observed $X$ and unobserved factors $\eta$ :

$$L^* = h(X) + \eta \tag{3}$$

Here, h is a non-parametric function of observed characteristics $X$.

$$L = max\{0, L^*\}, \; P(L^* < 0) > 0 \tag{4}$$

Finally, due to the presence of endogeneity, i.e., unobserved selection after controlling for observed characteristics, we can write the error term $\epsilon$ in Equation (2) as $\epsilon = \delta(X)$

$\eta + \varepsilon$, assuming $E[\varepsilon|L, X, \eta] = 0$, i.e., $\varepsilon$ is uncorrelated to $L$ conditional on observed and unobserved terms ($X$ and $\eta$) such that $\eta = L^* - h(X)$. Based on this, we can rewrite equation (1) as:

$$M = f(L, X; \beta) + g(X) + \delta(X)\eta + \varepsilon \tag{5}$$

This equation illustrates our first identification assumption (1) which we will call the linearity assumption, i.e., any unobserved confounding factor ($\epsilon$) can be represented as a linear function of $\eta$ such that the slope could change non-parametrically based on observed covariates $X$.

Using this framework, we leverage the phenomenon of bunching in mothers' labor supply $L$ to causally identify $\beta$. To understand this intuitively, consider mothers who have the same treatment and have bunched observations at zero hours of work $L = 0$. The remaining variation in this group of mothers can be controlled by the observed covariates $X$, hence, any systematic variation in $M$ among mothers with the same labor supply $L = 0$ conditional on $X$ must arise from variation in the unobservable confounders $\eta$ (or $L^*$) as seen from Equation (5). This enables us to build the control function by isolating the effect of $\eta$ on $M$ to identify $\beta$ as explained above.

To understand the details of how we build the control function, let's rewrite Equation (5) based on Equation (3) and Equation (4):

$$M = f(L, X; \beta) + g(X) + \delta(X)[\underbrace{L^* - h(X)}_{\eta}] + \varepsilon$$

$$M = f(L, X; \beta) + g(X) + \delta(X)[\underbrace{L + L^* \cdot 1(L = 0)}_{L^*} - h(X)] + \varepsilon$$

$$M = f(L, X; \beta) + \underbrace{g(X) - \delta(X)h(X)}_{m(X)} + \delta\,[L + L^* \cdot 1(L = 0)] + \varepsilon$$

$$M = f(L, X; \beta) + m(X) + \delta(X)\,[L + L^* \cdot 1(L = 0)] + \varepsilon \quad (*)$$

The model defined by Equations (3), (4), (5), and (*) implies:

$$E[M|L,X] = f(L,X;\beta) + m(X) + \delta(X)[L + E(L^*|L=0,X) \cdot 1(L=0)] \tag{6}$$

Note that for our homogeneous model, $\delta(X) = \delta$ and $f(L,X;\beta) = \beta$, s.t.

$$E[M|L,X] = \beta L + m(X) + \delta[L + E(L^*|L=0,X) \cdot 1(L=0)] \tag{7}$$

If we can identify the term $E[L^*|L=0,X]$ then, we can introduce the term $L + E[L^*|L=0,X] \cdot 1(L=0)$ into the equation as an additional control which will allow us to identify $\beta$ as well as $\delta$. $E[L^*|L=0,X]$ are the average type of mothers with $L=0$ and observed characteristics $X$, and as discussed earlier, we know that some of these mothers are of type $L^* < 0$, mothers who desire to work negative hours, i.e., are averse to working at all. However, it is difficult to ascertain how close these mothers are to being indifferent to working or not working, i.e., $L^* = 0$, and if we are able to get the average of these mothers, we can identify $\beta$. In order to identify this average type of mother, i.e., $E[L^*|L=0,X]$, we need to make additional distributional assumptions, beyond the linearity assumption as in Equation (5), about the distribution of $\eta|X$ which we call our second identification assumption. This assumption relies on distributional assumptions for $L^*$ given Equation (3). We will later show that our findings of $\beta$ estimates don't change even if we relax this assumption in the section on robustness checks.

Based on Caetano, Caetano, and Nielsen (2023), we follow three distribution assumptions:

*Assumption 2.1: Semi-parametric Normal*

$$\eta|X \sim \mathcal{N}(l(X), \sigma^2(X))$$

This does not require the assumption about the linearity of $m(X)$, $g(X)$, and $h(X)$, and about homoskedasticity.

*Assumption 2.2: Semi-parametric Uniform*

$$\eta|X \sim \mathcal{U}[\kappa(X), \mu(X)]$$

This assumption means that mothers at $L = 0$ have varied preferences for work, ranging from a minimum ($\kappa(X)$) to a maximum ($\mu(X)$) level, with an equal chance of falling within any specific range and none of them have $L^*$ values that are significantly far from a state of indifference.

*Assumption 2.3: Non-parametric Tail Symmetry*

*For all censored quantiles $q_0$, $\eta|X$ has symmetric tails below $q_0$ and above $1 - q_0$.*

This relaxes the normality assumption to ensure that extreme values of $\eta$ have an equal likelihood of occurring at both the lower (constrained) and upper parts of the distribution, making the distribution balanced at its extremes.

As outlined in Caetano et al. 2021, these assumptions are partially testable, and while normal distribution is a standard choice, we also explore other assumptions like non-parametric tail symmetry and uniform distributions to show the robustness of our results.

In summary, our empirical control function approach to identify causal impacts relies on two main assumptions: (1) the linearity assumption (selection-on-unobservables) in Equation (5), and (2): the assumptions on the distribution of $\eta|X$ which helps identify $E[L^*|L = 0, X]$ in Equation (6). Unlike traditional methods, this approach does not require the use of good instrumental variables which need exclusion restrictions to be satisfied and are difficult to find. Due to the non-negative constraint that mothers face, the average type of mothers with $L = 0$ and observed characteristics $X$ often have a negative desire to work, i.e., $E[L^*|L = 0, X]$ is often negative. Due to this, there is a discontinuity in the outcome for mothers at $L = 0$ which allows us to identify the treatment effects effect ($\beta$) by isolating the effect of the confounders ($\delta$) which are identifiable.

### 3.4 Discrete Clustering

In order to not solely depend on unobservable factors, while the observed control variables $X$ are used to serve as additional measures to account for endogenous variation, these do not play any special role in the identification strategy.

As discussed in Caetano, Caetano, and Nielsen (2023), to maintain the non-parametric nature of the model, it is recommended to "discretize" the covariates $X$ dividing them into distinct clusters denoted as $\{\hat{\mathcal{C}}_1, \ldots, \hat{\mathcal{C}}_K\}$, before we estimate $\mathbb{E}\left[L^* \mid L = 0, X\right]$. These clusters create a finite partition into the support of $X$, such that, $\hat{C}_K = (1(X \in \hat{\mathcal{C}}_1), \ldots, 1(X \in \hat{\mathcal{C}}_K))'$ is a vector of cluster indicators. Rather than using $X$ directly in estimating $\mathbb{E}\left[L^* \mid L = 0, X\right]$, we substitute $\hat{C}_K$ which has finite support due to the discretization. The estimator $\hat{\mathbb{E}}[L^* \mid L = 0, X] = \hat{\mathbb{E}}[L^* \mid L = 0, \hat{C}_K]$ is constructed using a two-step procedure: first, $X$ is discretized and second, one of the distributional assumptions is applied separately within each cluster. The clustering algorithm[1] is such that observations with similar values of $X$ are clustered together, and as the number of clusters ($K$) increases, the similarity in the values of $X$ within each cluster also increases due to a decrease in within-cluster variation in the controls. Hence, if $\mathbb{E}[L^* \mid L = 0, X]$ exhibits continuity, $\hat{\mathbb{E}}[L^* \mid L = 0, \hat{C}_K]$ will approximate this estimate more closely as $K$ expands.

To ensure that the function $m(X)$ in Equation (6) maintains a non-parametric structure as a function of $X$, we utilize the same clustering algorithm to define control where $m(X) = X'\tau + \sum_{k=1}^{K} \alpha_k 1\left(X \in \mathcal{C}_k\right)$. In this model, while the differences across clusters are controlled by cluster indicators non-parametrically, the linear controls account for differences within clusters attributed to $X$. As $K$ grows, the non-parametric match improves which reduces the unaccounted variation within each cluster progressively.

In summary, our empirical estimation follows the following steps:

[1] We use hierarchical clustering with the Gower measure of distance and Ward's linkage for its simplicity, stability, and ease of interpretation as we vary the number of clusters (Hastie, Tibshirani, and Friedman 2009).

(1) Group the data into $K$ clusters based on predetermined control variables X.

(2) For each cluster, we independently estimate $\mathbb{E}[L^* \mid L = 0, X \in \mathcal{C}_k]$, as explained above and construct the variable $(L + \mathbb{E}[L^* \mid L = 0, \hat{C_K}]1(L = 0))$.

(3) Incorporate this variable as an additional control in our main model.

In our homogeneous model, where $f(.)$ and $\delta(.)$ remain constant across different values of $X$, i.e., $f(L, X; \beta) = \beta L$ and $\delta(X) = \delta$, we estimate the model using the following Ordinary Least Squares (OLS) equation:

$$S = \beta L + X'\tau + \sum_{k=1}^{K} \alpha_k \mathbf{1}\left(X \in \hat{\mathcal{C}}_k\right) + \delta\left[L + \hat{\mathbb{E}}\left[L^* \mid L = 0, \hat{C}_K\right] \mathbf{1}(L = 0)\right] \tag{8}$$

# 4 Results

## 4.1 Homogeneous Impacts

In this section, we provide the estimates of $\beta$ for the specification of $f(.)$ based on the OLS model in Equation (8) under the different distributional assumptions about $\eta|X$ as discussed in the previous section. We also specify the $\delta$ estimates, i.e., the effect of the confounder $\eta$ on the mothers' mental health.

The results in Table 2 focus on our primary findings obtained from the entire sample. Column (i) presents the findings of the simple OLS regression of work hours on maternal mental health without any controls (Equation (8) without either $m(X)$ or the correction term). Column (ii) introduces the observed controls, $m(X)$ into the specifications from Column (i) (Equation (8) without the correction term). Finally, Columns (iii), (iv), and (v) present the corrected estimates of $\beta$ using our control function approach (Equation (8)) under the three different assumptions for the distribution of $L^*|X$ ($\eta|X$). Column (iii) assumes a uniform distribution (assumption 2.2), Column (iv) assumes normal distribution (assumption 2.1), and Column (v) assumes symmetry in the tails for distribution (assumption 2.3) for $L^*|X$, respectively. As discussed, we prefer the estimates under the standard assumption 1, however, we will show that our findings do not change drastically under the different distributional assumptions for the robustness of our results.

Table 2: Effect of Maternal Labor Supply on Mental Health

| | (i)<br>Uncorrected<br>No Controls | (ii)<br>Uncorrected<br>w/ Controls | (iii)<br>Semip.<br>Uniform | (iv)<br>Semip.<br>Normal | (v)<br>Nonp. Tail<br>Symmetric |
|---|---|---|---|---|---|
| $\beta$ | 0.045** | 0.039** | −0.293** | −0.274** | −0.379** |
| | (0.015) | (0.016) | (0.117) | (0.109) | (0.144) |
| $\delta$ | | | 0.280** | 0.259** | 0.363** |
| | | | (0.099) | (0.090) | (0.125) |

Notes: (1) The table shows estimates of the effect of an additional hour per week on maternal mental health score (MHI-5) for the full sample of mothers, $N = 5,866$. (2) Estimation using $K = 20$ clusters and $K_\delta = 1$. (3) Bootstrapped standard errors in parentheses (500 iterated samples). (4) The list of controls in $X$ specified in Section 2 . (***) $p < 0.01$, (**) $p < 0.05$, (*) $p < 0.1$.

In Table 2, finding from the simple naive regression without any controls, Column (i) shows that maternal mental health is strongly positively associated (at a 5% significance level) with their work hours, although this relationship becomes less positive after adding control variables, Column (ii). The discontinuity plots that are shown in Figure 4 suggest that these uncorrected estimates in Column (i) and Column (ii) are positively biased.

Results from the corrected estimates using the control function, Columns (iii)-(v), show that maternal labor supply has negatively significant effects on their mental health (MHI-5). Specifically, for every additional hour of work per week, the MHI-5 score, on average, decreases by 0.216 points based on *Assumption 1* (Semi-parametric Normal), Column (ii). The effect although small is still meaningful, especially when considering the cumulative effect of multiple hours over time. The $\delta$ estimates in the table show the average effect of the confounders $\eta$ (in terms of $L^*$) on the maternal mental health outcome, Column (iii) - (v). The results show a positively significant effect of these confounders suggesting mothers working more hours tend to be positively selected as compared to the ones who work lesser.

It is important to note what happens to our $\beta$ estimates as we progressively add control variables (selection-on-observables) and later correct for endogeneity using the selection-on-unobservables approach implemented by building a control function. Our estimates are positive without controls and reduced when observables are added and finally, the corrected estimates become negative. Even without looking at any significance levels, these results suggest the evidence that selection-on-unobservables matters for our causal estimation and that this selection is positive and in the same direction as the selection-on-observables. Also evident from Figure 4 (left panel), we see positively significant discontinuity in labor supply at $L = 0$ for the mental health score and as compared to $L > 0$, suggesting a positive effect of $L^*$ on $M$. The results from Figure 4 (right panel) show that even after controlling for all the observables, the discontinuity in the residualized mental health scores remains positive at $L = 0$ compared to scores at $L > 0$. This indicates that unobservables orthogonal to the observables add a positive bias to the main impact results from Column (ii).

# 5 Sensitivity and Robustness Analysis

In this section, we validate our key findings are robust by This section validates the robustness of the main homogeneous findings by relaxing the underlying assumptions of the control function approach outlined in Section 3.

Our empirical approach is based on two key assumptions. First, the linearity assumption, i.e., any unobserved confounding factor $\epsilon$ can be represented as a linear function of $\eta$ $(L^*)$ such that the slope could change non-parametrically based on observed covariates $X$ as in Equation (5). This restricts the effect of the confounders on the mental health outcome to be the same for $L = 0$ and $L > 0$. Second, the distributional assumption for the distribution of $L^*|X$ (assumptions 2.1, 2.2, 2.3).

We now will consider a violation of these assumptions. First, we will show that relaxing the distribution assumption does not affect our findings on the impact of labor supply on the mental health of mothers. Second, we assess the linearity assumption by employing two different approaches: (i) comparing local versus non-local extrapolation of the assumption; and (ii) exploring potential non-linearities, specifically variations in $\delta(X)$ by cluster.

Further, we check for the sensitivity of our results by changing the choice of the cluster count used in our control function specifications.

## 5.1 The Distributional Assumption

In this section, we show that our key findings are robust to the failure of the distributional assumptions about $L^*|X$. Specifically, we consider the variations in our estimates of $\beta$ that might occur when the estimator $\mathbb{E}[L^* \mid L=0, X]$ is biased due to the failure of the distributional assumptions. For this exercise, we denote $\tilde{E}$ as the expected value based on a specific distributional assumption and $\widetilde{\beta}$ as the resulting treatment effect computed through the control function method. Now, if the assumed distributional condition does not hold, $\tilde{E}$ may diverge from the actual $E$, which might lead to a biased $\widetilde{\beta}$.

To address this issue, I adopt the strategy outlined by Caetano, Caetano, and Nielsen (2023) and Caetano et al. (2021). The method involves expressing $\beta$ (accurate identification of true treatment effect) as a function of the misidentification of the expectation $E$:

$$B_\beta = -\frac{B_E}{\tilde{E}}\delta \tag{9}$$

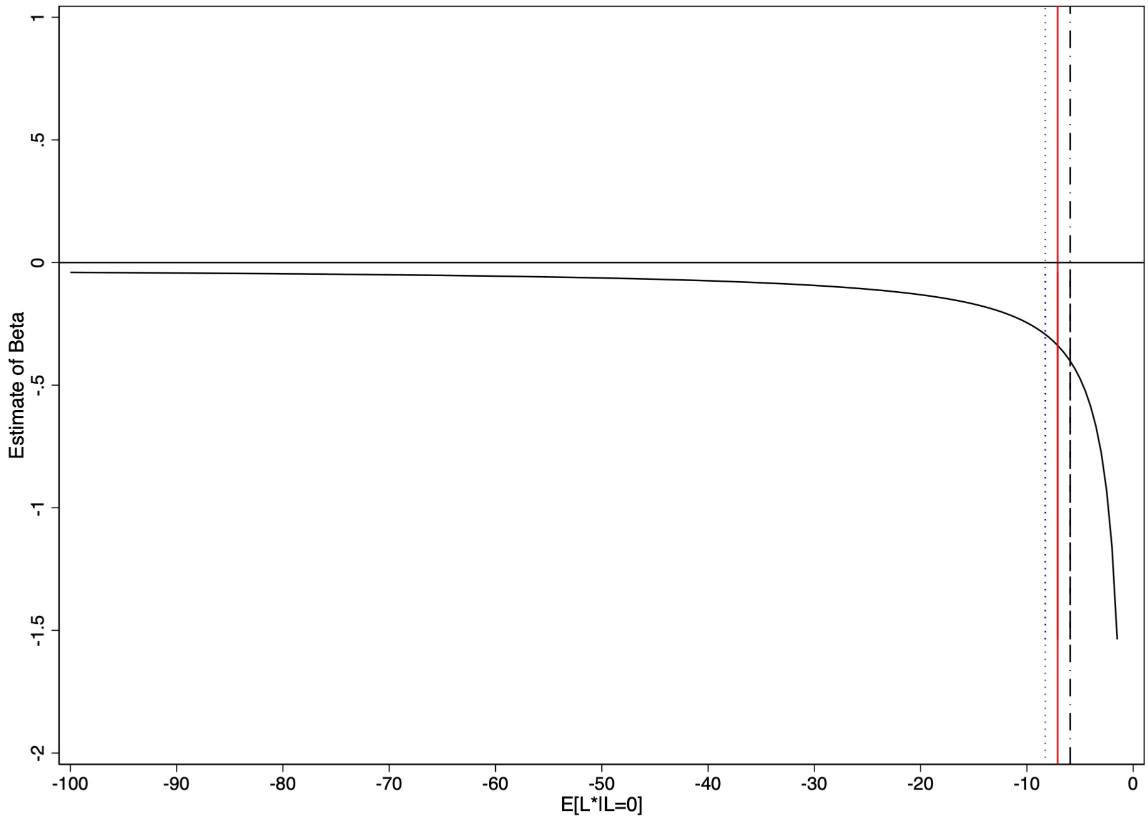
Here, $B_\beta$ signifies the bias in estimating $\beta$, and $\tilde{E}$ represents the biased expectation.

This formula reveals an asymmetry. For a specific value of $\delta$ and the magnitude of the expectation error $B_E$, it is preferable to have an estimator that inclines towards the expectation estimator $E(L^*|L = 0, X)$ when the expectation $E$ is highly negative rather than moderately negative. Figure 5 confirms this asymmetry. It illustrates the estimations of $\beta$ for various negative values of $\tilde{E}$ using a modified version of Equation (8):

$$S = \tilde{\beta}L + X'\tilde{\tau} + \sum_{k=1}^{K} \tilde{\alpha_k}\mathbf{1}\left(X \in \tilde{\mathcal{C}}_k\right) + \tilde{\delta}\left[L + \tilde{\mathbb{E}}\left[L^* \mid L = 0, \tilde{C}_K\right]\mathbf{1}(L = 0)\right] \tag{10}$$

The black curve in Figure 6 represents the calculated $\beta$ values that would have resulted from considering every potential mistaken expectation value. The vertical lines are included for reference and correspond to the weighted average expectation of $\hat{\mathbb{E}}\left[L^* \mid L = 0, X\right]$ estimated under the three different distributional assumptions described in Section 3 using 20 clusters. The estimates presented by the red solid line assume uniform distribution (Assumption 2.2), the blue dotted line assumes normality (Assumption 2.1), and the black dashed line assumes tail symmetry (Assumption 2.3). These correspond to expectations for our beta estimates in Table 2 in columns (iii), (iv), and (v), respectively.

Figure 6: Estimated $\beta$ for each counterfactual value of $\mathbb{E}[L^* \mid L = 0, X]$




Note: (1) The black curve shows the $\hat{\beta}$ that could be obtained from the regression of Equation (8) for different counterfactual values of $\tilde{E}$. (2) The vertical lines represent the weighted average of the estimates of $\mathbb{E}[L^* \mid L = 0, X]$ across all $K = 20$ clusters, obtained from the distributional assumptions: Semi-parametric Uniform (solid red line), Semi-parametric Normal (dotted blue line), and Non-parametric Tail Symmetry (dashed black line). (3) Number of observations: 5,866. Source: Author's calculation.

Our main takeaways are that maternal mental health is negatively affected by increasing their work hours ($\beta$) regardless of the distributional assumption we make since it does not depend on the value of $\tilde{E}$. Moreover, the estimate of $\beta$ shown in Table 2 (columns (iii)-(v)) may potentially be underestimated rather than overestimated because if we reported for more negative estimates of $\mathbb{E}[L^* \mid L = 0, X]$, the $\beta$ estimates would not change significantly under the distributional assumptions made. However, if we reported with less negative estimates of $\mathbb{E}[L^* \mid L = 0, X]$, $\beta$ estimates would become even more negative than our reported results. Therefore, while we cannot make any further inferences for the magnitude of our beta estimates, we can still conclude our findings would stay consistent qualitatively for any distributional assumption we make.

## 5.2 The Linearity Assumption

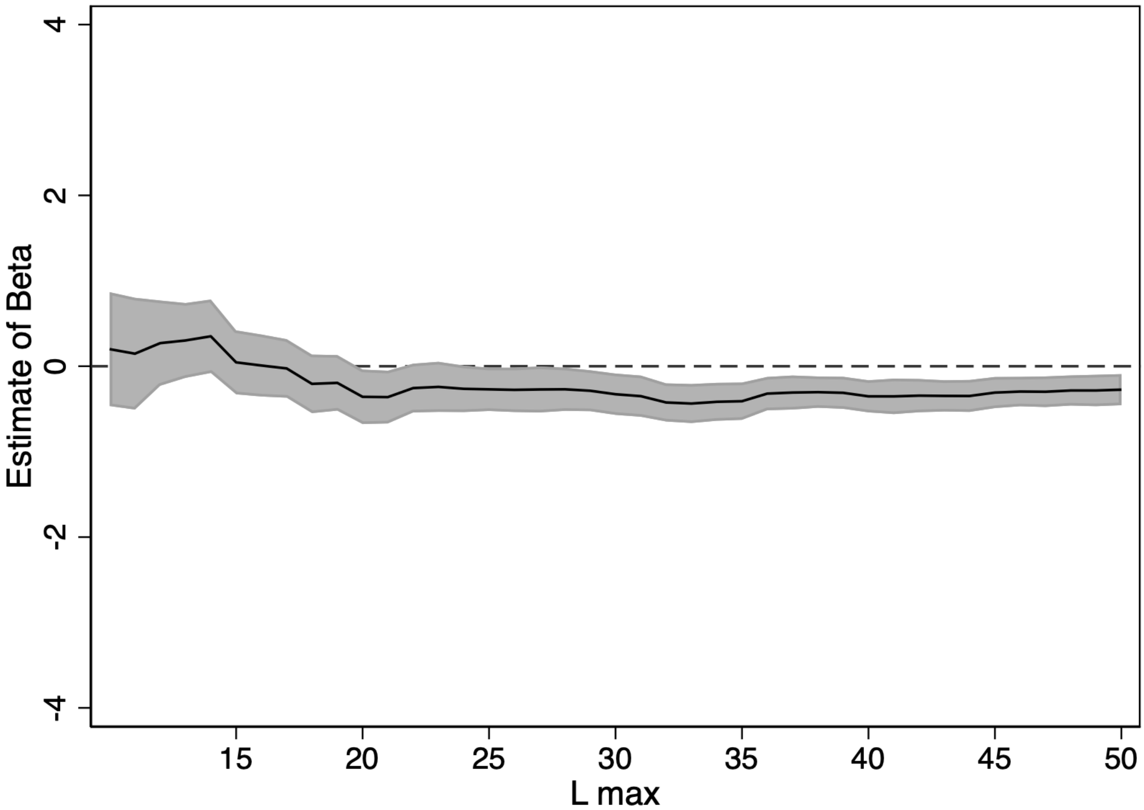
Our control function approach makes the linearity assumption which assumes confounder variables can represented as a linear combination of observed $X$ and unobserved factors $L^*$ which restricts the effect of $L^*$ on $M$ to be the same for $L = 0$ and $L > 0$. However, it may be possible that this effect differs for different levels of $L$. To explore this possibility, we perform two sets of sensitivity analyses to see how our main findings would change under violations of this assumption. First, we limit the degree of extrapolation of the effect of $L^*$ to smaller $L$ values, to see if our estimates change as the extrapolation degree extends towards larger $L$ values. Second, we introduce non-linearities in the effect of the confounder through $\delta$ to examine potential variations in our estimates across different $L$ values.

### 5.2.1 Local versus Non-local Extrapolation

For this first sensitivity analysis, we restrict the sample to $L \leq L_{max}$, incrementally increasing $L_{max}$ to make our global linearity assumption more local. Figure 7 shows the estimates of $\beta$ from the homogenous model (Table 2, column (v)) for different values of $L_{max}$. We should expect the beta estimates to differ as we relax the linearity assumption. However, overall, the truncated estimates are consistent with the full-sample results (far right estimates). Specifically, as $L_{max}$ gets closer to zero, the estimates get more imprecise due to a smaller sample size, however, as $L_{max}$ increases, the estimates are more precise and consistent with our results in Table 2, column (ii).

Despite the presence of noise for more localized estimates, we can still find evidence that our estimates are not significantly affected when we relax the linearity assumption and transition from a global to a local extrapolation.

Figure 7: Estimated $\beta$ for different samples with $L \leq L_{max}$




Note: (1) Figure plots the solid line estimates of $\beta$ for restricted samples with $L \leq L_{max}$. (2) Bootstrapped 95% confidence intervals based on 250 iterations shown in gray. (3) Estimates are shown for $L_{max} \geq 10$ and presented for Semi-parametric Normal distribution assumption. Source: Author's calculation.

### 5.2.2 Allowing for non-linearities

Another possible way to test the linearity assumption involves exploring the non-linear effects of the confounder $\eta$. For this exercise, instead of assuming a constant delta as in our homogeneous model case (Table 2, column (v)) results, i.e. $E[\epsilon|X, L^*] = m(X) + \delta L^*$, we can assume that it's plausible that $\delta$ varies non-parametrically with the cluster of the observed covariates $X$, i.e., $E[\epsilon|X, L^*] = m(X) + \delta(X)L^*$. This adjustment could account for some of the variation in $L^*$ across different $X$ values. This means we now are allowing confounders to have a different impact on the outcome for different values of $L$ indirectly, thus considering heterogeneity in the effects of the confounders around $L = 0$. If the confounding variables do exhibit non-linear effects, our beta estimated will be biased and these estimates will vary as we incorporate greater heterogeneity in $\delta(.)$ across $X$ clusters. Hence, we could modify our model specification with the $\delta(.)$ to allow for flexibility that would indirectly capture any non-linear effect of the confounders. Specifically, we define $\delta(X)$ as $\hat{C}_{K_d}(X)\delta$ such that $\hat{C}_{K_d}(X)$ represents a vector of indicators for each of the $K_d$ clusters of $X$, and $\delta$ is a $K_d$ dimensional vector. Table 3 demonstrates the robustness of results for this alternative specification where $\delta(X)$ can vary by clusters of $X$.

Table 3: Effect of Maternal Labor Supply on Mental Health: Allowing for $\delta(X)$ to vary by cluster

| | *(i)* Uncorrected No Controls | *(ii)* Uncorrected w/ Controls | *(iii)* Semip. Uniform | *(iv)* Semip. Normal | *(v)* Nonp. Tail Symmetric |
|---|---|---|---|---|---|
| $\beta$ | 0.045** | 0.039** | -0.367** | -0.314** | -0.434** |
| | (0.015) | (0.016) | (0.134) | (0.120) | (0.158) |
| $F(\delta)$ | | | 1.817 | 1.776 | 1.803 |
| | | | (0.920) | (0.928) | (0.920) |

Notes: (1) The table shows estimates of the effect of an additional hour per week on maternal mental health score (MHI-5) for the full sample of mothers, $N = 5,866$. (2) Estimation using $K = 20$ clusters and $K_\delta = 20$. (3) Bootstrapped standard errors in parentheses (500 iterated samples). (4) The list of controls in $X$ specified in Section 2 . (***) $p < 0.01$, (**) $p < 0.05$, (*) $p < 0.1$

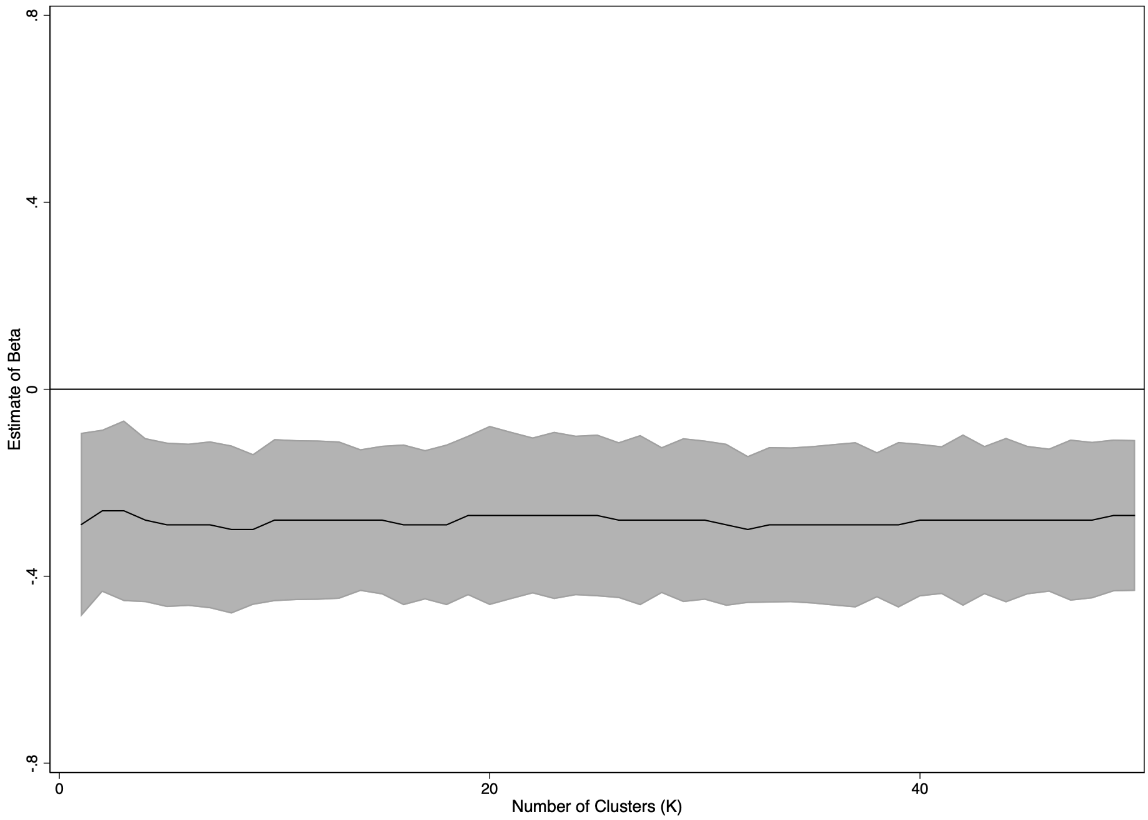
We find that our $\beta$ estimates as in Table 3, column (v) remain consistent with the previous findings after the change in specification, i.e., when we transition from a constant $\delta(.)$, i.e., $K_\delta = 1$ to a more unrestricted $\delta$, i.e., $K_\delta = 20$ to allow for heterogeneity across all 20 clusters of $X$. This indicates our findings are robust to permitting non-linearities in $L$ in our specifications.

In summary, we find that our estimated impacts seem to be robust to relaxing our linearity assumptions.

## 5.3 Cluster Choice: Controlling for $X$ flexibly

In this section, we examine whether our findings are robust to the choice we made regarding the number of clusters $K$ we use in our main specification results in Table 2, Columns (iii)-(v). The inclusion of clusters allows a non-parametric flexibility to $m(X)$ to account for a potential non-parametric effect of $X$ on $M$ at the cluster level. Additionally, clustering allows for variations in $\mathbb{E}[L^* \mid L = 0, X]$, i.e., the expectation may be the same within each cluster but can differ across clusters. As discussed in Section 3.4, as the number of clusters $K$ grows, the non-parametric match improves which reduces the unaccounted variation within each cluster progressively and helps to better approximate $m(X)$ as well as $\mathbb{E}[L^* \mid L = 0, X]$.

Figure 8: Estimated $\beta$ for different numbers of clusters $K$




Note: (1) The black line shows estimates of $\beta$ for different numbers of clusters $K = 1, 2, ..., 50$ and a 95% confidence interval. (2) Bootstraped standard errors using 250 bootstrap samples. (3) Estimates shown for distributional assumption Semi-parametric Tobit. (3) Number of observations: 5,866. Source: Author's calculation.

In Figure 8, we replicate our analysis with the preferred distributional assumption specification in Table 2, column (v) for different numbers of clusters $K = 1, 2, ...50$. As can be seen, results show that the impact estimate of maternal labor supply on their mental health scores remains consistently similar as $K$ increases. This suggests that our choice of 20 clusters in our specified models is sufficient for our approximation to account for the non-parametric structure of the functional forms and the expectations estimated.

# 6 Heterogeneity by Child Age and Part-time Work

We now explore the possible heterogeneous impacts of maternal labor supply on mental health based on the age of the youngest child and the mother's labor intensity. For the youngest child's age, we use an indicator variable for the mother having a child aged 5 years or less, and for the labor intensity, we use an indicator variable for the mother participating in part-time work (12 to 36 hours).

For the heterogeneity model, we re-define the functional form for $f(L, X; \beta)$ for the mental health outcome $M$ as followed by Caetano et al. (2021):

$$f(L, X; \beta) = (\beta + \beta_c \cdot C + \beta_p \cdot P + \beta_{pc} \cdot P \cdot C) \cdot L \tag{11}$$

Table 3 displays the results of the heterogeneity effect and we focus on the findings from our preferred distributional assumption of normality for $L^*|X$ in Column (iv) although results are similar across other distributions. Results reveal several noteworthy trends. Firstly, we find a negatively significant effect of labor supply on maternal mental health ($\beta < 0$) for mothers who don't have young children and are not working part-time. Secondly, the influence of work hours becomes larger ($\beta_c > 0$) or less negative for a mother who has a child less than equal to 5 years as hours worked increase, although the effect is statistically insignificant. Third, the impact of work hours on maternal mental health becomes more negative ($\beta_p > 0$) for mothers who participate in part-time work. This effect is statistically significant and implies that mothers who work part-time have better mental health. Finally, the effects of increasing work hours have a slightly more positive but insignificant impact ($\beta_{pc} > 0$) on mental health for mothers with young children who work part-time compared to mothers without young children who do not work part-time. This implies the negative impact of increased work hours on mental health is more pronounced for mothers with young children who work part-time.

For the $\delta(X)$ estimates, our results are still intuitive. While the overall positive selection in terms of the effect of confounders effect on mental health remains, we find a less positive selection in terms of mothers with young children which is insignificant but could mean the presence of a young child might provide emotional support and mitigate the negative impact of unobservables on maternal mental health.

In summary, the results suggest that increasing the labor supply negatively impacts maternal mental health, while part-time employment is associated with better mental well-being with mental health deteriorating for mothers who work part-time work if they increase their work hours. The presence of a young child doesn't significantly alter these relationships, indicating that part-time work remains advantageous for mothers, irrespective of the age of their child.

Table 4: Heterogeneity Effects of Maternal Labor Supply on Mental Health

| | *(i)* Uncorrected No Controls | *(ii)* Uncorrected w/ Controls | *(iii)* Semip. Uniform | *(iv)* Semip. Normal | *(v)* Nonp. Tail Symmetric |
|---|---|---|---|---|---|
| $\beta$ | 0.050** | 0.035 | -0.323** | -0.322** | -0.450** |
| | (0.019) | (0.021) | (0.144) | (0.131) | (0.177) |
| $\beta_c$ | -0.017 | 0.034 | 0.011 | 0.031 | 0.054 |
| | (0.029) | (0.037) | (0.236) | (0.229) | (0.288) |
| $\beta_p$ | -0.008 | -0.031 | -0.049** | -0.051** | -0.052** |
| | (0.022) | (0.022) | (0.023) | (0.023) | (0.023) |
| $\beta_{pc}$ | 0.024 | 0.053 | 0.049 | 0.051 | 0.052 |
| | (0.035) | (0.037) | (0.039) | (0.039) | (0.039) |
| $\delta$ | | | 0.313** | 0.305** | 0.433** |
| | | | (0.125) | (0.112) | (0.159) |
| $\delta_c$ | | | 0.013 | -0.002 | -0.023 |
| | | | (0.203) | (0.193) | (0.256) |

Notes: (1) The table shows estimates of the effect of an additional hour per week on maternal mental health score (MHI-5) for the full sample of mothers, $N = 5,866$. (2) Estimation using $K = 20$ clusters and $K_\delta = 1$. (3) Bootstrapped standard errors in parentheses (500 iterated samples). (4) The list of controls in $X$ specified in Section 2 . (***) $p < 0.01$, (**) $p < 0.05$, (*) $p < 0.1$.

# 7 Conclusions and Discussions

In this paper, we investigate the impact of maternal labor supply (weekly working hours) on their subjective well-being outcomes (MHI-5 scores). The analysis focuses on the Dutch population, utilizing data from the nationally representative LISS survey, CentER data in the Netherlands.

In order to address the possible endogeneity and selection bias, we employ a novel identification strategy of bunching estimation using a control function approach followed by Caetano, Caetano, and Nielsen (2023). This method corrects for endogeneity by leveraging the potential bunching of observations for the mothers at zero hours of work, a pattern arising from the non-negativity constraint on time. The argument is that mothers who have different confounding unobservables could still end up working the same, i.e., zero hours since they are at a corner-binding solution and cannot work negative hours (even if they want to). This helps to build a control function to isolate the average treatment effect of maternal labor supply by controlling the effect of confounders on mental health which is robust to different possible sources of endogeneity. This innovative method also allows us to look at heterogeneity impacts. Specifically, we explore the heterogeneity effects based on the age distribution of the youngest child and the intensity of labor quantity in terms of their participation in part-time work to understand the decisions mothers make about how much to work, especially for mothers with young children.

Our findings reveal a negative impact of increasing labor hours on maternal mental health, especially for mothers who participate in part-time work. Part-time employment is associated with better mental well-being with their mental health deteriorating if mothers who work part-time increase their work hours. The presence of a young child doesn't significantly alter these relationships, indicating that part-time work remains advantageous for mothers, irrespective of their children's age. We don't find our main results to change in case our identification assumptions do not hold true.

The mechanisms driving these effects could operate through several channels. These could include factors such as increased stress from balancing work and family responsibilities, greater workload demands within limited hours, limited workplace support, conflicting roles as caregivers and employees, social isolation, financial pressures, lack of autonomy in work schedules, and disrupted work-life balance. The institutional mechanisms supporting part-time work in the Netherlands can create an environment where mothers can effectively balance their work and family responsibilities. Specifically, in the Netherlands, part-time jobs have better quality relative to other OECD countries (Fagan et al, 2014), are much more institutionalized (Bosch, Deelen & Euwals, 2008), and have less involuntary work involvement rates (OECD, 2019). This could indicate that mothers are happy to work part-time and would corroborate our findings of deteriorating mental health for part-time working mothers if they increase their work hours by an additional hour. A deeper understanding of the different mechanisms driving our findings more quantitatively are areas of future research that we plan to explore further.

In conclusion, our study makes a valuable contribution, exploring and implementing a novel identification strategy to estimate the causal impact of maternal labor supply without the use of traditional methods which require instrumental variables with exclusion restrictions, randomized trials, or panel datasets for controlling fixed effects. Instead, we leverage the data which is bunched at zero work hours for mothers in our case to causally identify our $\beta$ estimates using a control function approach that relies on distributional assumptions about the confounder effects.

Our results also provide valuable conclusions for policy implications. These findings emphasize the importance of flexible work options, especially part-time employment, in supporting maternal mental health and overall family well-being, at least in the shorter run. The presence of these supportive mechanisms allows mothers to make effective choices on how much to work to enhance their overall quality of life and mental health. These policies could help mitigate the negative consequences of working longer hours on maternal subjective well-being and enable them to better balance their work and family responsibilities.

# 8 References

Allen, T. D., Herst, D. E., Bruck, C. S., & Sutton, M. (2000). Consequences associated with work-to-family conflict: a review and agenda for future research. Journal of occupational health psychology, 5(2), 278.

Balderson, U., Burchell, B., Kamerāde, D., Wang, S., & Coutts, A. (2021). An exploration of the multiple motivations for spending less time at work. Time & Society, 30(1), 55-77.

Bosch, N., Deelen, A., and Euwals, R. (2010). Is Part-time Employment Here to Stay? Working Hours of Dutch Women over Successive Generations. Labour, 24(1), 35-54.

Caetano, C. (2015). A test of exogeneity without instrumental variables in models with bunching. Econometrica 83(4), 1581–1600.

Caetano, C., Caetano, G., & Nielsen, E. (2023). Correcting for endogeneity in models with bunching. *Journal of Business & Economic Statistics*, 1-13.

Caetano, C., G. Caetano, E. Nielsen, and V. Sanfelice (2021). The effect of maternal labor supply on children: Evidence from bunching. Working Paper.

Campaña, J. C., Gimenez-Nadal, J. I., & Velilla, J. (2023). Measuring gender gaps in time allocation in europe. Social Indicators Research, 165(2), 519-553.

Dagher, R. K., Chen, J., and Thomas, S. B. (2015). Gender differences in mental health outcomes before, during, and after the Great Recession. PLOS ONE, 10(5), e0124103.

Ervin, J., Taouk, Y., Hewitt, B., & King, T. (2023). The association between unpaid labour and mental health in working-age adults in Australia from 2002 to 2020: a longitudinal population-based cohort study. The Lancet Public Health, 8(4), e276-e285.

Freeman, R.B. War of the Models: Which Labour Market Institutions for the 21st Century? Labor Economics. Vol. 5, 1998, pp. 1-24.

Fagan, C., Norman, H., Smith, M., and Menéndez, M. G. (2014). In search of good quality part-time employment.

Fogli, A. and L. Veldkamp (2011). Nature or nurture? learning and the geography of female labor force participation. Econometrica 79 (4), 1103–1138.

Greenhaus JH, Allen TD, Spector PE. Health consequences of work-family conflict: the dark side of the work-family interface. In: Perrewé PL, Ganster DC editors. Research in Occupational Stress and Well Being: Vol. 5. Employee Health, Coping and Methodologies. Elsevier Science/JAI Press (2006). p. 61–98.

Galván, E., & Garcıa-Penalosa, C. (2018). Gender norms and labour supply: Identifying heterogeneous patterns across groups of women. In Eighth Meeting of the Society for the Study of Economic Inequality.

Hammer, T. H., Saksvik, P. Ø., Nytrø, K., Torvatn, H., & Bayazit, M. (2004). Expanding the psychosocial work environment: workplace norms and work-family conflict as correlates of stress and health. Journal of occupational health psychology, 9(1), 83.

Higgins, C., Duxbury, L., and Johnson, K. L. (2000). Part-time work for women: Does it really help balance work and family? Human Resource Management: Published in Cooperation with the School of Business Administration, The University of Michigan and in alliance with the Society of Human Resources Management, 39(1), 17-32.

Khoudja, Y., & Fleischmann, F. (2018). Gender ideology and women's labor market transitions within couples in the Netherlands. Journal of Marriage and Family, 80(5), 1087-1106

Inanc, H. 2018. Unemployment, temporary work, and subjective well-being: the gendered effect of spousal labor market insecurity, American Sociological Review, vol. 83, no. 3, 536–66.

Layard, R. (2005). Rethinking public economics: The implications of rivalry and habit.

Economics and happiness, 1(1), 147-170.

Mache, S., Servaty, R., and Harth, V. (2020). Flexible work arrangements in open workspaces and relations to occupational stress, need for recovery and psychological detachment from work. Journal of Occupational Medicine and Toxicology, 15(1), 1-11

Minnotte, K. L., & Yucel, D. (2018). Work–family conflict, job insecurity, and health outcomes among US workers. Social Indicators Research, 139, 517-540.

Oishi, A. S., Chan, R. K., Wang, L. L. R., & Kim, J. H. (2015). Do part-time jobs mitigate workers' work–family conflict and enhance wellbeing? New evidence from four East-Asian societies. Social Indicators Research, 121(1), 5-25.

Pouwels, B., Siegers, J., & Vlasblom, J. D. (2008). Income, working hours, and happiness. Economics letters, 99(1), 72-74.

Piovani, C., & Aydiner-Avsar, N. (2021). Work time matters for mental health: A gender analysis of paid and unpaid labor in the United States. Review of Radical Political Economics, 53(4), 579-589.

Roxburgh, S. (2004): 'There Just Aren't Enough Hours in the Day': The Mental Health Consequences of Time Pressures. Journal of Health and Social Behavior, 32, pp. 115–31.

Rupert, P. A., Stevanovic, P., Hartman, E. R. T., Bryant, F. B., & Miller, A. (2012). Predicting work– family conflict and life satisfaction among professional psychologists. Professional Psychology: Research and Practice, 43(4), 341.

Sayer, L. C.; England, P.; Bittman & Bianchi, S. M. (2009): How long is the second (plus first) shift? Gender differences in paid, unpaid, and total work time in Australia and the United States. Journal of Comparative Family Studies, pp. 523-545.

Stoeva, A. Z., Chiu, R. K., and Greenhaus, J. H. (2002). Negative affectivity, role stress, and work–family conflict. J. Vocat. Behav. 60, 1–16. doi: 10.1007/s00420-014-0967-0.

Strandh, M., Hammarström, A., Nilsson, K., Nordenmark, M., & Russel, H. (2013). Unemployment, gender and mental health: the role of the gender regime. Sociology of Health & Illness, 35(5), 649-665.

Slotwinski, M., & Roth, A. (2020). Gender norms and income misreporting within households. ZEW-Centre for European Economic Research Discussion Paper, (20-001).

Thorsen, S. V., Rugulies, R., Hjarsbech, P. U., & Bjorner, J. B. (2013). The predictive value of mental health for long-term sickness absence: the Major Depression Inventory (MDI) and the Mental Health Inventory (MHI-5) compared. BMC medical research methodology, 13(1), 1-7.

Wang, S. and Coulter, R. 2019. Exploring ethnic and generational differences in gender role attitudes among immigrant populations in Britain: the role of neighborhood ethnic compos- ition, International Migration Review, vol. 53, no. 4, 1121–47

Warren, T. (2004). Working part-time: achieving a successful 'work-life'balance? The British journal of sociology, 55(1), 99-122.

Veenhoven, R. (2002). Why social policy needs subjective indicators. Social indicators research, 58, 33-46.

Visser, J. (2002). The first part-time economy in the world: a model to be followed? Journal of European Social Policy, 12(1), 23–42.

Zhang, Y., Duffy, J. F., & De Castillero, E. R. (2017). Do sleep disturbances mediate the association between work-family conflict and depressive symptoms among nurses? A cross-sectional study. Journal of psychiatric and mental health nursing, 24(8), 620-628.