

Spillover effects of long-term care insurance on spouse's labor supply

Xiao Han

Ph.D. candidate

Department of Risk Management and Insurance, School of Economics

Peking University

No.5 Yiheyuan Rd, Beijing, China, 100871

Tel: (86) 188 0011 5026

Email: hanxiaosky@pku.edu.cn

Zining Liu

Assistant professor

School of Insurance

Central University of Finance and Economics

Shahe District, Beijing, China, 102206

Tel: (86) 130 3113 5377

Email: ziningliu@cufe.edu.cn

Wei Zheng (*Corresponding author*)

Lloyd's Chair Professor

Department of Risk Management and Insurance, School of Economics

China Center for Insurance and Social Security Research (CCISSR)

Peking University

No.5 Yiheyuan Rd, Beijing, China, 100871

Tel: (8610) 6275 2058

Email: wzheng@pku.edu.cn

Spillover Effects of Long-term Care Insurance on Spouse's Labor Supply

By XIAO HAN, ZINING LIU, AND WEI ZHENG*

Expanding labor supply to sustain the economy will increasingly become a priority for China and other rapidly aging countries in the coming decades. Using data from China Health and Retirement Longitudinal Study (CHARLS), we provide a nationally representative estimate from China for the effect of long-term care insurance (LTCI) on spouses' labor supply. We exploit the introduction of public LTCI in 42 Chinese cities in a generalized difference-in-difference propensity score matching (PSM-DID) framework to explore the effect of LTCI and its mediation mechanism with careful attention to sample selection issues. Firstly, we find a significant and positive effect of LTCI on the labor supply of the insureds' spouses. This spillover effects exist both on the extensive margin (labor participation rate) and the intensive margin (prolonged working hours and rewards in earnings). Secondly, we find that the possible mediating mechanisms are spouses' potential health improvement and their re-allocation of time brought about by replacing informal family care with formal care services. Thirdly, we focus on the heterogeneous effects of LTCI on labor supply. We find that the effect varies by gender and age, and the LTCI policy design regarding target groups and reimbursement rules also significantly impacts the spillover effects. (JEL I18, J22, J31, I10)

I. Introduction

For people with chronic illnesses or disabilities in a rapidly aging population, long-term care (LTC) services are in great demand (Feng et al., 2020). For instance, across 26 European OECD countries in 2017, 50% of people aged 65 and over (65+) reported having at least some limitations in their daily activities, among which 17% reported severe limitations (Indicators and Hagvísar, 2019). Some studies expect that on average, the publicly funded LTCI expenditure in OECD countries would reach 2.5% of GDP in 2060 (Oliveira and Maisonneuve, 2015). The increasing number of aging

* Han: Department of Risk Management and Insurance, School of Economics, Peking University, No.5 Yiheyuan Rd, Beijing, China, 100871 (e-mail: hanxiaosky@pku.edu.cn); Liu: School of Insurance, Central University of Finance and Economics, Shahe District, Beijing, China, 102206 (e-mail: ziningliu@cufe.edu.cn); Zheng: Department of Risk Management and Insurance, School of Economics, Peking University, and China Center for Insurance and Social Security Research (CCISSR), Peking University, No.5 Yiheyuan Rd, Beijing, China, 100871 (e-mail: wzheng@pku.edu.cn).

disabled also raises the demand for informal family care. Previous literature shows that informal care is preferred over institutional formal care by care recipients and their relatives (Mentzakis et al., 2009; Lipszyc et al., 2012; Blaise, 2018; Hajek et al., 2018). For elderly disabled, their spouses are often one of the most common type of care providers (especially wives), who are predicted to be an increasingly important source of informal family care considering the declining fertility rate, the decreasing number of children, and the expanding elderly population (Ryan et al., 2012).

Meanwhile, aging is also increasingly exerting pressure on the labor market. Governments in different countries are trying to prolong people's working lives and increase their labor market participation (Fischer and Müller, 2020). Under the current demographic changes, more and more people stay in the labor force market until their late sixties.¹ Given that most caregivers are in their age of 50+ (Johnson and Wiener, 2006; Van Houtven et al., 2013; Schmitz and Westphal, 2017), they are faced with a time conflict between labor supply and informal care provision (Fischer and Müller, 2020). Analysis of labor supply among the elderly (e.g., spouses of the disabled) will help inform family and labor policy formulations in the context of an aging society.

In response to the mounting demand for LTC, many countries have launched long-term care insurance (LTCI) systems, aiming at socializing the responsibility of LTC for the elderly, considering their rising life expectancy (Chen et al., 2020), shrinking household size (Carmichael and Charles, 2003), and the increasing number of working wives (Liu et al., 2014). As an essential component in the modern social security system, LTCI mainly covers the daily care expenses or provides additional financial support for individuals with disabilities caused by illness or injury. Previous studies mainly focus on the direct effects of LTCI on health status (Jeon and Kwon, 2017), labor supply (Fu et al., 2017), consumption (Ariizumi, 2008), and investment decisions (Davidoff, 2010) of the disabled themselves. However, health shocks may affect not just the individuals who experience them but also their family members as a whole (Jeon and Pohl, 2017). Studies have noticed that formal care service presents a substitute for informal family care (Charles and Sevak, 2005; Stabile et al., 2006; Cremer et al., 2016), and LTCI is thus expected to have spillover effects on family caregivers by reducing their economic burden, physical toil and psychological

¹ Statistics in Japan and the US reveal labor participation rates at 42.7% and 30.8% for people aged 65-69 years old to participate in the labor force market (Statistics, 2016; Japan, 2016).

pressure and increasing their labor supply accordingly (Kumamoto et al., 2006; Bobinac et al., 2010; Levine et al., 2010; Arai and Zarit, 2011; Bakx et al., 2015). Family caregivers may choose to increase their probability of labor participation or re-enter the labor market (Fu et al., 2017) after getting rid of the heavy burden of care. The assessment of caregivers' response to LTCI is of great significance to offer a more comprehensive insight into the evaluation of LTCI arrangements from the perspective of family welfare and the social economy. This paper contributes several ways to the inconclusive existing studies on the spillover effects of LTCI on caregivers' labor supply among individuals with heterogeneous characteristics.

Firstly, we try to answer the question whether LTCI has a significant positive effect on caregivers' labor supply. Studies on caregivers' labor market outcomes of LTCI are relatively scarce. To the best of our knowledge, there are very few studies on the spillover effects of LTCI on caregivers' labor supply in a narrow sense², and their results are controversial. Fu et al. (2017) found a significantly increased labor participation rate of caregivers after the introduction of LTCI in Japan in 2000 and a decrease in labor supply after the amendment of LTCI in 2006, during which the benefit of LTCI was reduced. Kim and Lim (2015) found that the crowd-out effect of Korean LTCI on informal care was heterogeneous among caregivers who provided services to the disabled elderly at different health levels. Geyer and Korfhage (2015a; 2015b; 2017) found that benefit in cash of German LTCI inhibited family caregivers' labor supply, indicating that not all LTCI could stimulate the labor supply of family caregivers. To shed light on such effects, we analyze how one's availability to LTCI affects his/her spouse's employment and earnings and provide new mechanical and empirical evidence of LTCI's effect on family caregivers' labor supply, using China as a case study. We also investigate the LTCI's heterogeneous spillover effects on labor supply across different groups.

Secondly, this paper is one of the first to look at the effect of LTCI on caregivers' labor market performance from multiple dimensions. Previous studies show that

² Although there are plenty of studies exploring the causal impact of informal care activities on labor supply and earnings (Wolf and Soldo, 1994; Ettner, 1995; Ettner, 1996; Carmichael and Charles, 1998; Johnson and Sasso, 2000; Dentinger and Clarkberg, 2002; Carmichael and Charles, 2003; McGarry, 2007; Heitmueller, 2007; Bracke et al., 2008; Bolin et al., 2008; Lilly et al., 2010; Casado-Marín et al., 2011; Ciani, 2012; Van Houtven et al., 2013; Sugawara and Nakamura, 2014; Crespo and Mira, 2014; Heger, 2014; Yamada and Shimizutani, 2015; Fukahori et al., 2015; Skira, 2015; Schmitz and Westphal, 2017; Niimi, 2018), literature exploring the effect of long-term care insurance on caregivers' labor supply is indeed quite few.

formal care usually conducts as a substitute to family informal care (Mellor, 2001; Charles and Sevak, 2005; Stabile et al., 2006; Kim and Lim, 2015). Public LTCI often provides generous formal care services and/or subsidies, which is expected to mitigate the time conflict between caregivers' labor supply and informal care provision. We focus on the simultaneous influence of LTCI on caregivers' career choices and working behaviors on the extensive and intensive margin. First of all, we investigate the extensive margin whether economically active caregivers opt to increase their probability in labor participation when their spouses are under the coverage of LTCI, employing a difference-in-difference propensity score matching (PSM-DID) method. Then, we investigate the intensive margin that whether caregivers increase their total working hours and obtain more hourly income caused by various possible factors, including the improvement of productivity, taking on more responsibilities, and pursuing a workplace promotion (Sageer et al., 2012; Jaskiewicz and Tulenko, 2012). We also study caregivers' job-hopping behavior within and between self-employed and employed work, the responsibilities, working conditions, and time schedules of which are quite different by nature (Schmitz, 2016; Hassan, 2016).

Thirdly, this paper provides causal evidence about how the effect of LTCI transits to caregivers' labor supply. We identify this causal effect through two mediating indicators: caregivers' health status and their re-allocation of time. Literature shows that higher levels of health are associated with more time spent in the labor market (Ozturk and Kose, 2019). For caregivers, informal care burden can harm their physical health (Do et al., 2015) and mental health (Coe and Van Houtven, 2009), making caregivers away from the competitive labor market. LTCI's provision of accessible formal care services may reduce the care burden and improve caregivers' health level (Schmitz and Westphal, 2015), which further increases labor supply. In this paper, we analyze whether LTCI could promote spouses' labor market performance by preventing potential health impairments. Meanwhile, rational individuals also face the choice of allocating time to labor supply, enjoyment of leisure, and provision of informal care (Hallberg, 2003; Heitmueller, 2007; Bredtmann and Vonnahme, 2019). Evidence on the substitution and complementary relationship between informal care and leisure time are both found. Some studies argue that caregivers cut back on leisure time in order to provide informal care (McGarry, 2006), while others argue that individuals reduce their labor supply to provide care to their sick spouses and enjoy joint leisure with their family at the same

time (Michaud and Vermeulen, 2011; Jeon and Pohl, 2017) under their hearty patient-centered values (Chiang and Wu, 2020). The final effect of LTCI on leisure time is an open question, leading to the indecisive effects on the spillover effects of LTCI on caregivers' total working hours. In this paper, we analyze whether LTCI could promote spouses' labor market performance by reducing leisure time.

Results confirm that informal care provision puts restrictions on caregivers' labor supply and LTCI stimulates their willingness to work by reducing their care burden. We find a significant increase in spouses' labor participation rate (+5.9%) upon the introduction of LTCI. Total working hours rise by about 80.78 hours/year (or 13.3 minutes/day) and the hourly salary of paid employment rises by about 2.76 RMB yuan (approximately \$0.43). Affected caregivers' increases in labor supply are driven mainly by promoted health level and reduced informal care hours and joint leisure time. In heterogeneity analysis, we find that after the introduction of LTCI, women's increase in labor supply is much more than that of men, indicating a larger tendency to become a caregiver for wives than husbands. We also find evidence that young spouses are more willing to participate in work and pursue a promising career than older ones. Moreover, caregivers in LTCI pilot cities with both severely and moderately disabled individuals covered are found to provide more labor force compared to those in cities with only severely disabled individuals covered, mainly attributed to the weakening substitution relationship between formal and informal care among more severely disabled persons. High-intensity-caregivers for totally care-dependent individuals seem to give no response to LTCI and may choose to accompany their beloved care recipients, who are suffering from severe illness and at a risk of death (Fischer and Müller, 2020). Caregivers are also unaffected in LTCI pilot cities where the insurance only provides cash transfers instead of subsidies of formal care services. These findings hold under various robustness checks and are absent in placebo regressions.

The remainder of this paper is organized as follows. In Section II, we give detailed information about the institutional features of LTCI in China and draw our hypotheses from former literature. Section III describes the dataset and the construction of variables. Section IV presents the empirical framework, including methods of PSM-DID and the mediating analysis. Section V shows the results. Section VI provides a heterogeneous analysis. Section VII gives several robustness checks. Section VIII concludes and provides a brief discussion.

II. Background and Literature Review

A. Long-term Care Insurance Initiatives in China

To assure the elderly with affordable care services, in June 2016, public long-term care insurance (LTCI) pilots was launched in China with the release of the official document “Guiding Opinions on the Pilot Program of Long-term Care Insurance System” (Human Resources and Social Security Ministry of China, 2016). This document announced the first batch of 15 pilot cities. In September 2020, the central government decided to expand the LTCI pilots to additional 14 cities. Besides the above pilots, some other cities started their own LTCI programs. Until April 2021, 77 cities in 27 provinces have issued detailed plans of LTCI (see Table 1 and Fig. 1). All of the pilots are financed by existing public health insurance programs — the Urban Employee Basic Medical Insurance (UEBMI) that covers formal-sector employees and the Urban-Rural Resident Basic Medical Insurance (URRBMI)³ that covers other residents in urban and rural areas. LTCI fund is a fixed amount or a particular percentage of the existing public health insurance funds. Some pilots also have supplemental funding from the government (and welfare lotteries), individuals, or employers. There is a significant variation in funding standards, ranging from 30 to 700 RMB yuan (around \$4.6 to \$107.7) each year per person. All 77 cities cover UEBMI enrollees, while only a handful also cover URRBMI enrollees. Individuals in pilot cities with the corresponding public health insurance participate in LTCI automatically.

Typical practices in some pilots are summarized in Table A.1 in Appendix A. We also summarized critical features of LTCI pilot programs, including fundraising channels and benefits of different LTC services in Table A.2 and Table A.3. Eligibility for LTCI benefit is stringent (see Fig. A.1 in Appendix A). In most cities, beneficiaries must be severely disabled for at least six months, evaluated by assessing their ability to carry out activities of daily living (ADL) based on the Barthel score or its variant. Two categories of services in three locations are covered: nursing care and daily living care in hospitals/institutions/at home. Eligible insureds can apply for both

³ Including urban resident medical insurance and new rural cooperative medical insurance in the old system, and urban-rural resident basic medical insurance in the new system.

services in one of the above locations according to their severity of disability. Home care services are provided at beneficiaries' residences or health centers in local communities, including basic life care (such as bathing and turning over), common clinical care (such as nasal feeding and catheterization), risk prevention guidance (such as fall prevention guidance), and functional maintenance/rehabilitation training (such as eating training, passive joint activities). Services of institutional care are often provided in aged care facilities, including long-term residence, meals, necessities of daily lives, and also basic life care, primary clinical care, risk prevention guidance, and functional maintenance/rehabilitation training. Hospital care services are usually provided in local hospitals, including long-term residence and more medical-intensive clinical care. Benefits packages vary from city to city. Some pilots reimburse users with a fixed percentage of the total care expenditures with a cap. Others reimburse a fixed amount per day (or month) with limits on the total service hours (or days). To encourage home care provided by relatives, beneficiaries in some pilots who received home care continuously over a specific length of time are provided with extra cash benefits.

In previous literature, effects of these pilots on the access, quality, and cost of LTC have been evaluated (Yang et al., 2016; Wang et al., 2018; Lu et al., 2017; Li, 2018; Yang et al., 2018; Zhu and Österle, 2019; Zhang and Yu, 2019; Ma et al., 2019; Chang et al., 2020) with quite different findings partly attributed to their distinctive institutional backgrounds in different cities. Under the cooperation of local governments and private insurance companies, the LTCI system is now well-organized in many pilots, although details of LTCI policy rules are still undergoing adjustments to ensure an appropriate and favorable institutional design. Variations in LTCI programs enable us to investigate how policy features affect caregivers' labor supply in different regions across China.

[Insert Table 1 here]

[Insert Figure 1 here]

B. Literature Review and Research Hypotheses

Previous literature indicates that informal family care is the primary source of care services for the disabled in the absence of a convenient and affordable formal care system (Meng and Tang, 2010; Bhattacharyya et al., 2011; Lu et al., 2017). Studies also show that family caregivers are providing informal care at the expense of their personal health (Schmitz and Westphal, 2015) and reduced working hours (Michaud and Van Soest, 2008; Coe and Van Houtven, 2009). Many studies have noticed a substitution relationship between informal care provided by family caregivers and the formal care provided by hospitals and nursing homes (Charles and Sevak, 2005; Stabile et al., 2006; Cremer et al., 2016; Huang and Fu, 2017). Therefore, formal care provided and subsidized by LTCI may crowd out informal family care and thus effectively reduce family caregivers' burden (Kumamoto et al., 2006; Arai and Zarit, 2011), which may improve caregivers' health status and lead to a further increase in labor supply.

In our paper, we focus on three important issues related to the effect of LTCI on labor supply, which were not investigated by previous studies. The first one is the inconclusive results in related literature about the impact of LTCI on the labor market for both theoretical research and empirical research (Lilly et al., 2007). On the one hand, LTCI demonstrates an "opportunity cost effect" on family caregivers' labor supply. As a way of self-insurance, after adverse health shocks, family caregivers will reduce labor supply and spend more time accompanying and providing informal care for the disabled family member (Bobinac et al., 2010; Jeon and Pohl, 2017). Benefits provided by LTCI reduces the marginal cost of formal care (equal to the opportunity cost of informal care), thereby reducing family caregivers' incentives to provide informal care, which will increase their labor supply accordingly (Kim and Lim, 2015; Geyer et al., 2017; Fu et al., 2017). On the other hand, LTCI has a "consumption smoothing effect". Health risks usually have a negative impact on family annual income (Blundell et al., 2007; Blundell et al., 2016). After the health shock, the

spouse and other dependents' financial situation is invariably affected if the primary earner of the family loses ability to generate income. To smooth household's consumption under different health states and guarantee a basic standard of living, family caregivers will increase labor supply to make up for lost income, which is called the "added worker effect" (AWE) (Lundberg, 1985; Stephens, 2002). LTCI benefit protects family members from severe poverty and weakens the "added worker effect", resulting in a reduction in family caregivers' labor supply⁴. Therefore, the net impact of LTCI on labor supply is inconclusive and determined by these two opposite effects.

[Insert Figure 2 here]

Previous studies focus on the effect of LTCI on informal care (Kim and Lim, 2015) or the effect of informal care on caregivers' labor supply at both the extensive margin and the intensive margin (Lilly et al., 2007; Schmitz and Westphal, 2017). Other studies focus on informal care's effect on labor market performance via family caregivers' physical and mental health (Bauer and Sousa-Poza, 2015; Kohl et al., 2019; Rellstab et al., 2020). Fu et al. (2017) investigate the effect of LTCI on labor supply just at the extensive margin while without exploring the effect on the labor market performance. To our knowledge, there is limited research about the effect of LTCI on the labor market performance of caregivers. Furthermore, previous studies indicate that the continuous disposable time released by LTCI increases family caregivers' probability of choosing full-time jobs (Blundell, Pistaferri, et al., 2016) and thus improves their working productivity by a richer workplace experience as well as increases their wage rate accordingly. In this context, we try to fill the gap and analyze whether LTCI works effectively in improving the labor market performance of caregivers from multiple dimensions.

Regarding the controversy in existing research, we try to figure out the net spillover effects of LTCI on caregivers' labor market performance and come up with the following Hypothesis 1. We will explore it in section V.

⁴ However, Chu et al. (2018) found that when faced with health shocks, medical insurance payments and other extra guarantees did not promptly become a risk compensation tool for rural households to smooth consumption. Family members still need to increase labor supply when the effect of LTCI is difficult to effectively demonstrate in a short time.

Hypothesis 1: LTCI improves caregivers' labor market performance (including the extensive and intensive margins).

The second issue in the related literature is how LTCI affects labor supply via mediation factors. Studies found that informal family care would increase caregivers' physical burden and mental pressure, reduce their social interaction levels (Sugisawa et al., 1997), and hurt their physical and mental health (Michaud and Van Soest, 2008; Coe and Van Houtven, 2009; Zhang et al., 2009). Moreover, family caregivers are constrained by the time length of informal care, and they usually have insufficient time to seek medical treatment and cannot recover from the onerous duties in time (Yu and Feng, 2018). The deteriorated health resulting from informal care may have two adverse effects on caregivers' labor market performance. On the one hand, the deteriorated health will lead to a decline in labor productivity, resulting in obstacles to getting employed or promoted. On the other hand, individuals in poor conditions tend to quit the labor market (Li and Tan, 2019). Studies found that unhealthy physical and mental conditions significantly reduced their productivity and hourly income, which further reduced their enthusiasm in labor participation (Rellstab et al., 2020). Therefore, informal care has negative effect on labor supply via deteriorated health status, while LTCI can reduce informal care burden and improve the labor market performance via improving health status (Kumamoto et al., 2006; Arai and Zarit, 2011; Strumpf, 2011). In this paper, we investigate the role of health status as a mediating factor influencing the spillover effects of LTCI on caregivers' labor supply. This leads to our Hypothesis 2.1, which will be explored it in section V.

[Insert Figure 3 here]

Moreover, LTCI may affect labor supply via changing time allocation of different time-consuming activities. To be specific, individuals need to choose different time-consuming activities under certain time endowments. These activities include working, leisure, and provision of informal care in our research. Some studies argue that caregivers often cut back on leisure time and labor supply in order to provide informal care (White-Means and Chang, 1994; McGarry, 2006). Other studies argue that individuals only reduce their labor supply to provide informal care to their sick spouses and enjoy joint leisure at the same time (Michaud and Vermeulen, 2011; Jeon

and Pohl, 2017). Caregivers also tend to seek more leisure activities to relieve their psychological pressure and painful experience in caregiving (Schüz et al., 2015). After the introduction of LTCI, the provision of convenient and cheap formal care services could release more disposable time for caregivers to re-allocate. The effect of LTCI on leisure time is inconclusive, making the effect of LTCI on total working hours indecisive. This paper will explore this time re-allocation effect of LTCI on family caregivers and figure out their labor supply decisions based on their different socio-demographic characteristics. The redistribution of time between leisure activities and labor supply is also affected by individuals' preferences and other external factors (Aguar et al., 2021). We try to provide empirical evidence for the mediation mechanism of leisure activities on LTCI's spillover effects of caregivers' labor supply. This leads to our Hypothesis 2.2, which will be explored it in section V.

[Insert Figure 4 here]

Hypothesis 2.1 (mediation effect of health level): LTCI affects caregivers' labor supply via physical and mental health.

Hypothesis 2.2 (mediation effect of time re-allocation): LTCI affects caregivers' labor supply via caring time and leisure time.

The third issue in related literature is the heterogeneous effect of LTCI on labor supply across different groups and different LTCI policies. Considering the differences in the labor supply elasticity of individuals with different characteristics, there may exist a discrepancy in the spillover effects of LTCI on caregivers' labor supply across different groups by gender, age, income, living arrangement, education level, and retirement status (Carmichael and Charles, 2003; Blundell et al., 2016; Fu et al., 2017). We are particularly concerned about the heterogeneity of the LTCI spillover effects in different groups by gender and age. For the age issue, as the life expectancy becomes longer in many countries, more and more elderly decide to leave the labor market later (Hirazawa and Yakita, 2017). However, individuals of different ages have diversified labor market responses and labor supply elasticity to policy changes, and the response of older retirees is quite modest because of their relatively low labor capability and low elasticity (Lee et al., 2016). For the gender issue,

previous studies believe that women are more likely to become caregivers than men, and LTCI may thus have a more significant impact on wives' labor supply than husbands. In the US, for instance, female caregivers are more likely to provide intensive care for their parents, resulting in a lower probability of labor participation than non-caregivers (Skira, 2015). In Canada, the labor participation rate of female caregivers was also found lower than that of non-caregivers (Lilly et al., 2010), and in Japan, studies have also shown that informal care provided by female caregivers has a negative impact on their labor supply (Sugawara and Nakamura, 2014; Yamada and Shimizutani, 2015). This paper explores the heterogeneous responses in labor supply among caregivers for different groups by age and gender. Here we develop Hypothesis 3.1 and will explore it in section VI.

Moreover, we try to figure out whether the eligibility for LTCI application would affect the its spillover effects. Previous studies find that formal care is not necessarily a substitute for informal family care. As the level of disability of the elderly person increases, the substitution relationship between informal care and formal care tends to disappear. It even demonstrates a gradually stronger complementary relationship in the particular period before the death of the elderly (Bonsang, 2009; Huang and Fu, 2017). These results highlight the heterogeneous effects of informal care on formal care and suggest that informal care is an adequate substitute for formal care only if the care needs of the elderly are low and require unskilled types of care. In some LTCI pilot cities, both severely disabled (with a Barthel score less than 40) and moderately disabled (with a Barthel score more than 40 and less than 60) elderly can apply for benefits; but in other pilots, only severely disabled elderly is eligible LTCI benefits. The impact of LTCI on the labor supply of caregivers for moderately and severely disabled people is expected to be different.

Finally, we are concerned about whether benefit in kind or benefit in cash would promote caregivers' labor supply in a more efficient way. Studies show that LTCI benefit in cash may decrease family caregivers' labor supply. Geyer et al., (2017) found that the introduction of German LTCI which provided benefits in cash had a negative impact on men caregivers' labor supply but had no significant impact on women caregivers. We develop Hypothesis 3.2 and explore the heterogeneous spillover effects of LTCI on caregivers considering different eligibility and benefits of LTCI policies in section VI.

Hypothesis 3.1 (heterogeneity of the spillover effects): The spillover effects of LTCI vary across different groups by age and gender.

Hypothesis 3.2 (heterogeneity of the spillover effects): The spillover effects of LTCI vary across different LTCI policies regarding eligibility and benefits.

III. Data and Variables

A. Data

We use China Health and Retirement Longitudinal Study (CHARLS) to investigate the effect of LTCI on individuals' labor supply whose husband/wife was recently covered by LTCI. CHARLS belongs to a family of well-established international health and retirement datasets including, for example, the Health and Retirement Study (HRS) in the US, the English Longitudinal Study of Ageing (ELSA) in England, and the Survey of Health, Aging, and Retirement in Europe (SHARE). CHARLS surveys nationwide residents mainly aged 45 years and older (with no upper age limit) and their spouses, totaling 17.7 thousand individuals living in 10.3 thousand households in 450 villages/communities (Zhao et al., 2014) in each survey year. A stratified (by per capita GDP of urban districts and rural counties) multi-stage PPS (probabilities proportional to size) random sampling strategy was adopted. Following detailed protocols for sampling, field survey, and data quality verification, the CHARLS national baseline survey was firstly conducted in 2011 in 150 counties of 28 provinces across China. The second-, third- and fourth-wave national surveys in 2013, 2015, and 2018 aimed to revisit the same respondents sampled in the first wave. Information in CHARLS includes respondents' demographic backgrounds, family features, health levels, cognition and depression, health care and medical insurance⁵, working hours and corresponding income (income and costs of self-employed work, and wage and benefits of employed work), retirement and pension, expenditures and assets, housing property, and basic community conditions. CHARLS is also one of the best choices for studies of aging issues because of its focus on senior people, with more than 97.2% of respondents 45 years of age or older. To analyze the dynamic changes of individual's labor participation rate in different years and compare the

⁵ Questions about LTCI were newly introduced into CHARLS in 2018. However, most of the interviewees did not answer these questions.

working hours and hourly income for different groups, we select the sample of individuals who were continuously interviewed in 2011, 2013, 2015, and 2018 by ruling out respondents with missing information in one of the four years. We keep the variables consistently measured across the data set and finally obtain a balanced four-year panel data with 13,019 respondents in each year.

B. Treatment Group, Control Group, and “Intention-to-Treat Effect on the Treated”

Individuals with UEBMI and (or) URRBMI in pilot cities participate in LTCI automatically, providing a quasi-natural experimental setting in which we can examine the impact of LTCI on their spouses’ labor supply. We could determine whether an individual is covered by LTCI by figuring out whether his/her city of residence belongs to LTCI pilots and whether the local LTCI program is eligible within their basic health insurance. From Table 1, we find that only 42 pilot cities⁶ launched their own LTCI programs before July 2018, after which the last wave of CHARLS was conducted.⁷

[Insert Figure 5 here]

Note that we do not observe directly whether an individual actually received LTCI benefits. We can only observe the city of his/her residence but have no information concerning other eligibility criteria (e.g., bedridden for more than six months due to illness, disability, etc., and being assessed for LTC needs with a Barthel score less than 40). Here, we estimate an “intention-to-treat effect on the treated (ITT)⁸” (Dolls et al., 2018) instead of the traditional “average treatment effect on the treated (ATT)”. We examine the overall policy effect on the targeted population (all individuals in pilot cities who meet LTCI enrollment conditions) regardless of whether he/she has already received LTCI benefits. Therefore, ITT analysis tends to underestimate the

⁶ CHARLS only covers 26 of the 42 LTCI pilot cities: Qingdao, Weifang, Shangrao, Jinan, Beijing, Jilin, Jingmen, Chengde, Shanghai, Anqing, Hangzhou, Chengdu, Xuzhou, Guangzhou, Linyi, Jiaying, Linfen, Qiqihar, Suzhou, Liaocheng, Ningbo, Chongqing, Binzhou, Taizhou, Changsha, and Weihai. However, there is no information in CHARLS about the other 16 LTCI pilot cities: Changchun, Rizhao, Nantong, Xingtai, Songyuan, Shihezi, Tonghua, Tai’an, Baishan, Jining, Karamy, Zibo, Changji, Dongying, Heze, and Yantai.

⁷ CHARLS surveys were generally carried out in August during each investigation year, so we focus on cities where the LTCI policy was conducted before July 2018.

⁸ ITT analysis examines the average treatment effect on the targeted population of LTCI policy instead of the population actually affected by the policy.

traditional ATT effect, making our results more convincing if we still drew a significant conclusion under ITT estimation. The policy effect estimated here (ITT) is the product of ATT (average treatment effect on individuals who actually receive LTCI benefit) and the proportion of individuals who actually receive LTCI benefit (Fisher et al., 1990; Gupta, 2011). In CHARLS, the proportion of individuals with moderate and severe disability is about 17%, making the actual estimation of the spillover effects about six times larger than our results.

$$ITT = ATT \times \frac{\text{Individuals who actually receive LTCI benefit}}{\text{All individuals in the pilot cities who meet the conditions of LTCI enrollment}}$$

Under our research framework, the treatment group includes individuals whose spouses reside in pilot cities that launched the LTCI program before July 2018 and are covered by corresponding basic medical insurance⁹. The control group includes individuals whose spouses reside in cities without LTCI or in cities with LTCI but are not covered by corresponding basic medical insurance. The LTCI policy variable $Treatment_i$ is defined as a dummy variable ($Treatment_i = 1$ for the treatment group, $Treatment_i = 0$ for the control group).

C. Outcome Variables and Covariates

We could clarify the spillover effects of LTCI through changes in outcome variables before and after the policy shock. In the main analysis, we first examine the effect of LTCI on the labor participation rate of spouses and then extend our analysis to both their working hours and hourly income. We confirm each spouse’s current status of work by a question in the “Work, Retirement and Pension” questionnaire: “Did you engage in XXX work for at least 10 days in the past year?” and consider individuals to participate in the labor market if they self-report “yes”, or currently do not work but self-report looking for work. We consider individuals “not working” if they self-report that they currently do not work or have retired.

Next, we focus on the spillover effects of LTCI on the intensive margin of labor supply. Previous studies indicate that there are considerable differences in the recruitment process, working schedule, salary payment method and level, and glass ceilings between individuals employed by others and by themselves, resulting in

⁹ It should be noted that the treatment group covered by LTCI vary in different pilot areas. All pilots cover UEBMI enrollees, while only a handful also cover URRBMI enrollees.

different elasticity of working hours and hourly income with these jobs (Barón and Cobb-Clark, 2010; Noon et al., 2013; Mandel and Semyonov, 2014). Therefore, we analyze the working hours and hourly income of three types of work. Specifically, we studied the aggregated time of different types of labor in the last year, including 1) yearly hours on self-employed work, 2) yearly hours on employed work, and 3) total yearly hours of the both. Moreover, we studied individuals' corresponding hourly income of different types of work in the past year, including 1) hourly income of self-employed work, 2) hourly income of employed work, and 3) average hourly income of the both. To clarify, we regard informal family care as “care time”, which is mutually exclusive with “labor time” or “leisure time”.

Furthermore, we focus on two mediation mechanisms of the spillover effects of LTCI: health levels of caregivers and their re-allocation of time. For the first mechanism, the mediating variables are: 1) number of chronic diseases¹⁰, and 2) total score for depression¹¹, which are widely used as typical health factors affecting labor supply (Zhang et al., 2009; Langley et al., 2010; Ojeda et al., 2010; Garthwaite et al., 2014). For the second mechanism, the relevant variables are: 1) yearly hours of informal family care, and 2) yearly hours of leisure. Covariates include variables that have an influence on individuals' labor supply, including individual characteristics (age, gender, ethnicity, marital status, education level, self-rated health, outpatient frequency, inpatient frequency, medical insurance participation¹²) and household characteristics (family size, number of children, number of siblings, whether a house owner, whether living alone, annual family income, deposit, and financial assets). Outcome variables and covariates in the basic and mediating analysis are listed in Table 2, and descriptive statistics are listed in Table 3.

¹⁰ The number of chronic diseases (e.g., heart problems) of caregivers themselves has a significant negative effect on their willingness to provide informal care and their work productivity (Ågren et al., 2010). Researchers also find that individuals who are suffering inconvenience and chronic pains from diseases report a lower subjective health level and lower satisfaction concerning daily lives (Ganapathy et al., 2015).

¹¹ Yikilkan et al. (2014) find a negative effect of one's depression level and psychological illness on caregivers' labor supply.

¹² The related question is: “Are you currently covered by the following medical insurance?” 1. Urban employee medical insurance (yi-bao); 2. Urban resident medical insurance; 3. New cooperative medical insurance (he-zuo-yi-liao); 4. Urban and rural resident medical insurance; 5. Government medical insurance (gong-fei); 6. Medical aid; 7. Private medical insurance: purchased by work unit; 8. Private medical insurance: purchased by an individual; 9. Urban non-employed person's health insurance; 10. Other medical insurance.

[Insert Table 2 here]

[Insert Table 3 here]

IV. Empirical Strategies

A. PSM-DID Method: Does LTCI Affect the Labor Supply of Spouses?

There is potential endogeneity between labor supply and caregiving activities where researchers argue that the informal caregiving activities stem from their lower labor supply (Lilly et al., 2007). Previous studies suggested that certain family members have a low tendency to participate in the labor market, leading them to be more willing to provide informal care instead of applying for LTCI benefits (Henz, 2004; Van Houtven et al., 2013). Some studies solve the endogenous problem using the instrumental variable (IV) method (Bolin et al., 2008; Coe and Van Houtven, 2009; Heitmueller, 2007; Leigh, 2010; Van Houtven et al., 2013). However, such studies did not take the possibility that the labor supply trend before the LTCI introduction will affect informal care provision into consideration, leaving the potential endogenous problems still unsolved (Fu et al., 2017). To deal with this problem, we use the difference-in-difference propensity score matching (PSM-DID) method to reduce the influence of selection bias and solve the endogeneity problem caused by the omission of covariates and nonlinear terms (Caliendo and Kopeinig, 2008). This method has been used in many pieces of literature to study the impact of LTCI on medical expenditure (Ma et al., 2019) and family caregivers' labor supply (Geyer and Korfhage, 2015a).

The essence of PSM-DID method is to compare the change in outcome variables Y_t of treated individuals before and after the introduction of LTCI $E(Y_{t+s}^1 - Y_t^0 | LTCI = 1)$ to a counterfactual change $E(Y_{t+s}^0 - Y_t^0 | LTCI = 1)$, as if they were not affected by the policy changes. The counterfactual change, in turn, is an actual change for controlled individuals, i.e., $E(Y_{t+s}^0 - Y_t^0 | LTCI = 0)$ if a common trend assumption is satisfied: $E(Y_{t+s}^0 - Y_t^0 | LTCI = 1) = E(Y_{t+s}^0 - Y_t^0 | LTCI = 0)$. The assumption means that without the intervention of LTCI policies, the trend of labor supply in the treatment group and the control group should be the same. From 2011 to 2018, LTCI pilots have been launched in different cities one after another, so we apply a multi-period PSM-DID method to estimate the spillover effects with the following regression equation,

[Insert Formula 1 here]

Where $Labor_supply_{it}$ represents the outcome variables and $Year_t$ is a series of year dummy variables with its coefficient β_1 representing the year fixed effect. α_i

is a series of individual dummy variables and its coefficient β_2 measures the individual fixed effect that does not change over time. $Treatment_i$ is a dummy variable for distinguishing the treatment and control groups. $Time_{it}$ is also a dummy variable that marks the introduction of LTCI. Individuals in different cities participate LTCI at different time. For our four-period data, $Time_{it} = (0,1,1,1)$, $Time_{it} = (0,0,1,1)$ or $Time_{it} = (0,0,0,1)$ indicates that individuals began to be covered by LTCI at the second (2013)¹³, the third (2015)¹⁴ or the fourth (2018)¹⁵ period, respectively. The coefficient β_3 of $Time_{it} * Treatment_i$ represents the change in outcome variables of the treatment group relative to the control group after the introduction of LTCI, indicating the policy effect of LTCI. X_{it} is a vector composed of a series of covariates. For randomly assigned treatment group, the parallel trends assumption should be satisfied,

[Insert Formula 2 here]

Where $j = -3, -2, -1$ represents the third, second, and the first period before the introduction of LTCI in different pilot cities. Its coefficient γ_j represents the difference in outcome variables between the treatment and control groups before LTCI was first introduced to the city. However, the random assignment of treatment and control groups, or the common trend assumption, is a crucial criterion for DID method. In practice, researchers often utilize a similar criterion to determine the two groups by nature or by uncontrollable factors (J. Heckman et al., 1998). When we assess the influence of LTCI on labor supply, ideally, the control group should be individuals randomly or naturally extracted to be unaffected by LTCI. Unfortunately, such an assumption is often difficult to satisfy. There are non-negligible differences in economic development, disability severity of the elderly, and family care mode between cities of treatment and control groups. Before designing and implementing LTCI policies, the government usually takes regional conditions into account, including the socio-economic level and the adequacy of medical insurance funds of a city, making the choice of a pilot far from a natural experiment. It is also unethical to select old persons randomly to receive LTCI benefits.

To make the common trend assumption more credible, alternatively, we extract the controlled respondents to match the treated ones based on their similarity in terms of

¹³ Qingdao city.

¹⁴ Weifang city and Changchun city.

¹⁵ Other 39 pilot cities.

propensity scores. The propensity score $p(LTCI = 1 | X_{it})$ measures the probability of being treated conditional on a set of covariates X_{it} , where only controlled individuals with propensity scores overlapped within a bandwidth to those treated are extracted for DID estimations. PSM-DID can remove the selection bias and confounding bias caused by the non-randomness of the selection of LTCI pilot (Rosenbaum and Rubin, 1983) and eliminate the unobservable individual heterogeneity that does not change over time (Bryson et al., 2002). We first derive the propensity score using logistic regression:

[Insert Formula 3 here]

The above equation predicts the probability of being treated and defines it as the propensity score. X_{it} is a set of covariates representing respondents' individual and household characteristics. According to preceding studies, individuals who are female, elderly, with a smaller family size, and lack substitutes for informal care are more likely to become caregivers and less likely to enter the labor force. People with unfavorable health status (e.g., poor self-rated health) may also self-selected into caregiving (Schmitz and Westphal, 2015) and stay out of the labor force market. We take these factors into consideration for our PSM analysis (Austin, 2011; Brookhart et al., 2006; Garrido et al., 2014).

We then match and compare each treated respondent's outcome variables to a weighted set of outcome variables of controlled respondents through the Gaussian kernel matching method. Specifically, each individual in the control group will be weighted inversely by their distance in propensity score from that of treated individuals within a bandwidth (i.e., 0.05) of the score (Heckman et al., 1997). Individuals with a propensity score located outside the bandwidth will be disregarded from the common support domain. We then check whether Rubin's $B < 25$ and whether Rubin's $R \in (0.5, 2)$ in order to pass the balancing test (Rubin, 2001) and make sure that the matched treatment and control group is randomly assigned. We then could trace the changes in ITT after the introduction of LTCI to obtain the PSM-DID estimator as follow,

[Insert Formula 4 here]

Here the subscript i and j denote individuals in treatment group T and control group C , respectively; S represents the area of common support, and w_{ij} shows the corresponding weight for kernel matching. The estimators ITT catch the improvement

of individuals' labor supply with the introduction of LTCI after removing systematic differences between treated and controlled individuals. The advantage of this method lies in that it can make full use of the sample information. It is widely used in the evaluation of policy and reform effects (Zhou and Cao, 2017).

B. Mediation Analysis: Are Health Level and Time Re-allocation Mediating Variables of the Spillover Effects of LTCI?

After figuring out the spillover effects of LTCI on individual's labor supply, we are also interested in the mediating effect of LTCI on individual's labor supply in which the mediating variable is health level and time re-allocation. We use $Channel_{it}$ to denote the observed health level or leisure time. Because the health level (leisure time) can be affected by LTCI, there exist two potential values, $Channel_{it}^1$ and $Channel_{it}^0$, only one of which will be observed. For example, if individual i actually covered by LTCI, then we observe $Channel_{it}^1$ but not $Channel_{it}^0$. We use $Labor_{it}(Channel_{it})$ to denote the outcome variable, labor supply, resulted from the mediating variable $Channel_{it}$. The causal mediation effects or indirect effects for each individual i as follows:

[Insert Formula 5 here]

The above equation examines the change of the outcome variables Y_{it} if the mediators $Channel_{it}$ change from the value of individuals without LTCI (namely $Channel_{it}^0$), to the value of individuals with LTCI (namely $Channel_{it}^1$), which could be observed in the real world. In section V, we will apply the Sequential Regression Coefficient method of Baron and Kenny (1986) and use the KHB method (Karlson and Holm, 2011; Karlson et al., 2010; Breen et al., 2013) to decompose the total effect into direct and indirect effects. The KHB method is a general decomposition method that is unaffected by the rescaling or attenuation bias that arises in cross-model comparisons in nonlinear models. It recovers the degree to which a covariate, $Channel_{it}$, mediates or explains the relationship between independent variables and a latent outcome variable, underlying the probability model (Kohler et al., 2011). Firstly, we examine whether LTCI can significantly affect the mediation factors $Channel_{it}$ of spouses: their leisure time or health level. Secondly, we explore whether there exists a significant correlation between LTCI and spouses' labor supply without the control of mediation factors $Channel_{it}$. Finally, we analyze the impact of LTCI on labor supply

with the covariate $Channel_{it}$ and obtain the policy effect after ruling out the mediation effect. If the policy effect of LTCI on labor supply in the third step smaller than the one in the second step, we believe that the mediation effect exists.

Step 1: [Insert Formula 6 here]

Step 2: [Insert Formula 7 here]

Step 3: [Insert Formula 8 here]

The corresponding hypothesis are as follows:

[Insert Formula 9 here]

If H_0^1 , H_0^2 and H_0^3 could be rejected at the same time, we will conclude that there exists significant mediation effect. Moreover, if $H_0^4 : \zeta = 0$ could also be rejected, then we regard $Channel_{it}$ as an incomplete mediation effect. However, if $H_0^4 : \zeta = 0$ could not be rejected, then we call $Channel_{it}$ a complete mediation effect. According to the Difference of Coefficient Methods (MacKinnon et al., 2007), we can obtain the numerically estimate of mediation effect by computing $\eta - \zeta$ in this linear case. Similarly, the Product of Coefficients Method regards $\kappa\gamma$ as a valid estimate of the causal mediation effect (MacKinnon et al., 2002) as long as the Sequential Ignorability Assumption (“SIA”; Imai et al., 2010) and the Linearity Assumption holds (Jo, 2008). SIA confirms that the selection of the LTCI pilot is random and independent of the mediation variables $Channel_{it}$ and the outcome variable $Labor_supply_{it}$, and that $Channel_{it}$ satisfies the exogenous assumption. SIA could be met after we conduct a propensity score matching and obtain two groups of individuals with similar characteristics except for their probability of being treated (Pan and Bai, 2015). Under another framework, we also examine whether $H_0^5 : \kappa\gamma = 0$ could be rejected (Sobel, 1982). If $H_0^5 : \kappa\gamma = 0$ is rejected, we believe there exists a mediation effect and we will further identify the proportion of medication effects and direct effects in the total effect (MacKinnon et al., 2002). The relationship between mediation effect (or indirect effect), direct effect and total effect is $Total_effect = Direct_effect + Mediation_effect$.

V. Results

A. *The Spillover Effects of LTCI on Spouses' Labor Supply*

We show the sample size, mean and median standardized differences (SD) across all covariates before and after kernel propensity score matching (Table 4). Comparing the size of matched subsamples on common support to that of unmatched ones, we confirm that over 98% of the treated individuals have at least one controlled respondent staying within the 0.05 bandwidth of their propensity scores.

[Insert Table 4 here]

The mean and median standardized differences across all covariates provide a framework to test the extent to which the treated individuals are quasi-experimentally randomized. Randomization will be achieved when treated respondents are matched with highly similar controls such that the mean and median SD shall be trivial between treatment and control groups. We find that the matched samples all passed the balancing test, with Rubin's $B = 14.7 < 25$ and Rubin's $R = 0.88 \in (0.5, 2)$, indicating the common trend assumption between treatment and control groups is met. However, the overall balances in all covariates, nonetheless, do not guarantee balance in each covariate. We further test the balance in mean SD of each covariate in the post-PSM subsamples. Although there is no generally agreed criterion upon significant imbalances of covariates, maximum SD at 10% is taken to signify negligible differences in previous studies (Austin, 2011). We find in Table 5 that in the post-matching subsample, all the covariates involved pass the balancing test with the absolute standardized bias of the covariates less than 10%.

[Insert Table 5 here]

[Insert Figure 6 here]

Fig. 6 shows the results of tests for parallel trends after propensity score matching. It can be found that the outcome variables and potential mediation factors (individuals' physical health, mental health, and leisure hours per year) all passed the parallel trend tests in 2011, 2013, and 2015. However, in 2018, some variables demonstrate a coefficient significantly different from zero. Fig. 6 shows that after the propensity score matching, the core outcome variables and mediating variables all passed the parallel trend test.

A.1 Extensive Margin: Probability of Labor Participation

This section examines the effect of LTCI on spouses' labor participation. We first examine the matched individuals' total probability of labor participation (TPLP) and the probability of participating in self-employed work (SPLP), and paid employment (EPLP). The results are shown in Table 6.

[Insert Table 6 here]

From Table 6, we find that LTCI does improve the labor participation rate of individuals whose husbands/wives reside in LTCI pilot cities. Specifically, after the introduction of LTCI, spouses become 5.9% more likely to enter the labor force market than individuals whose husbands/wives reside outside the range of LTCI benefits. Furthermore, treated spouses in LTCI pilot cities are 4.0% more likely to participate in self-employed work, and a slightly more likely to participate in paid employment (2.0%, insignificant). On the one hand, attributed to the flexible time schedule of self-employed work, the corresponding labor participation rate could be adjusted immediately after the introduction of LTCI and increased 4.0% than the control group. On the other hand, things are different with jobs of paid employment. The searching of paid employment jobs often takes a longer time and involves more complicated process, so it is difficult for individuals to get employed in a relatively short time period.

Regarding the types of work chosen by family caregivers, Yamada and Shimizutani (2015) found that the impact of informal care on labor supply differs in terms of working intensity of a job, and Dispenza et al. (2019) found that living with a disabled family member affects people's career choice. Self-employed work is widely chosen among caregivers partly because it allows individuals to arrange time schedules according to their preference and obtain more discretionary time (Thompson and Prottas, 2006). However, we examine the spillover impact of LTCI on spouses' willingness to participate in self-employed non-agricultural work and find a negative effect (not shown in Table 6).¹⁶ In fact, according to the answers of CHARLS respondents, individuals usually choose private or family business work mainly because of their own unsatisfactory health reasons, the need for caring for family members, or being forced to leave. The greater pressure associated with ownership of a small business detracts from the advantages of having autonomy (Thompson and Prottas, 2006). After the introduction of LTCI, spouses who had to engage in overloaded work in order to maintain fundamental living expenses might stop their self-employed non-agricultural jobs.

¹⁶ Due to space limitations, the results of the impact of LTCI on the subcategories of self-employed work (self-employed agricultural work and self-employed non-agricultural work) have not been shown in the article. Readers can request from the author.

A.2 Intensive Margin: Yearly Working Hours

In recent studies, Bolin et al. (2008) and Van Houtven et al. (2013) measure the impact of family caregiving activities on not only the extensive margin (i.e., labor participation) but also the intensive margin (i.e., working hours and their hourly income). We estimate the spillover effects of LTCI on the total and subdivided kinds of working hours in Table 7.

We find that the total yearly hours of work increased significantly after the introduction of LTCI. Specifically, LTCI plays a role of helper for family caregivers in pilot cities, releasing about 81 hours of annual labor supply for each spouse (or 13.3 minutes/day). Besides, LTCI resulted in a 96-hour's significant increase in self-employed working hours per year. Since self-employed work is often relatively flexible and can be adjusted conveniently and promptly, individuals tend to prolong their hours devoted to these types of work immediately after reducing informal care hours. By contrast, the time schedule of paid employment is often fixed during a particular period. It is more difficult for employees to extend their working hours in a short period. In Table 7, we find a smaller and significant positive spillover effect (67 hours per year) on paid employment.

[Insert Table 7 here]

A.3 Intensive Margin: Hourly Income

Many studies have explored informal care's impact on caregivers' hourly salaries (Schmitz and Westphal, 2017). Reduced care burden by LTCI may have a positive impact on their own physical and mental health level, which can stimulate the enthusiasm of family caregivers to seek a promising career and promote their productivity in the workplace (Bauer and Sousa-Poza, 2015; Kohl et al., 2019; Rellstab et al., 2020). Furthermore, previous studies find a positive correlation between working hours per day and working productivity (Blundell et al., 2016), indicating better workplace performance of a full-time job than a part-time job. LTCI may increase caregivers' working hours and health status, further improving their productivity and wage rate. We estimate their yearly working hours and hourly income in Table 8.

In Table 8, we do not find a significant effect of LTCI on spouses' average hourly income and the hourly income of self-employed work. By contrast, we find a positive effect of LTCI introduction on hourly income of employed work. Individuals whose spouses are covered and subsidized by LTCI have a higher hourly wage (2.76 RMB yuan/hour or \$0.43/hour) than those without LTCI. Combined with the results above, we find that although the probability of labor participation may not be increased obviously in a short time for paid employment, their hourly wages indeed increased significantly, indicating a potential higher productivity. In general, the improvement of caregivers' working enthusiasm is reflected differently across labor categories. LTCI increases working hours for self-employed workers, while for employed workers, LTCI improves their working hours and productivity simultaneously.

[Insert Table 8 here]

B. Causal Mediation Effects

In causal mediation analysis, we examine the indirect effects of LTCI on the outcome variables (labor supply) through mediation factors (Robins and Greenland, 1992). As mentioned above, there may exist two channels concerning the effect of LTCI on spouses' labor supply. Firstly, LTCI can reduce spouses' care burden and improve their physical and mental health, and thus promote their labor supply both on the extensive and intensive margins. Secondly, the benefit and reimbursement of LTCI make the formal care services more cost-effective and will reduce the provision of its substitute, informal care. The joint leisure time with family members may also be reduced at the same time to provide more working hours and make more money. Thereby, family members may have more disposable time to re-allocate and increase their labor supply.

B.1 Mechanism I: Spillover Effects Through the Improved Health Level of Spouses

Empirical results indicate that a higher health level is often associated with more time spent in the labor market (Ozturk and Kose, 2019). Informal care burden can harm caregivers' physical and mental health (Barnay and Juin, 2016), increase the number of chronic diseases and the depression risk of caregivers, making them away from the labor market. LTCI's provision of accessible formal care services can reduce such burden and thereby improve individuals' health and their enthusiasm in labor supply. In our following analysis, we choose two variables as mediation indicators about individuals' physical and mental health: the number of chronic diseases and the depression score. Table 9 shows the health promotion effect of LTCI.

[Insert Table 9 here]

Firstly, it can be seen from Table 9 that LTCI significantly reduced the number of spouses' chronic diseases (-0.161, significantly), indicating that people's physical health has indeed been improved after LTCI's introduction. It may result in a potential increase in their willingness to work (Ganapathy et al., 2015). Secondly, the heavy care burden for the disabled may also harm the mental status of caregivers (Yıkılkan et al., 2014). Previous studies have shown that informal care significantly increases the risk of depression for caregivers (Alfakhri et al., 2018). Fortunately, LTCI reduces caregivers' degree of depression by a reduction in depression score (-0.095).

Next, we examine whether LTCI increases spouses' hours of labor market performance by improving his/her physical and mental health. In many countries, chronic disease is the major cause of death and disability and a reflection of significant changes in dietary habits, physical activity levels, and tobacco use (Halpin et al., 2010; Waxman, 2004). Individuals with chronic diseases are usually faced with significantly reduced productivity at work, unemployment, and premature retirement (Zhang et al., 2009). After the provision of more convenient and cheaper formal care services, spouses of the disabled could may be able to escape from the heavy burden and get access to a healthier lifestyle. Besides, previous studies find that mental illness is associated with a significantly poorer labor market performance (Chatterji et al., 2007; Andersen, 2015). LTCI improved spouses' mental state by reducing their depression score and would thus increase their willingness to work and their work performance.

We examine whether the effects of LTCI on spouses' labor supply become smaller or insignificant after adding the variable of physical health level (measured by the number of chronic diseases) to the covariates. Table 10 confirms this mediating mechanism, from which we find a smaller or a decreased significant level after we controlled for spouses' physical health level. Moreover, we conduct PSM-DID analysis in groups with different physical health levels. We find from Table 11 that spouses in a better physical health condition are more likely to participate in the labor market than those in a relatively poorer physical health condition (7.7% significant compared with 3.8% insignificant). We also witness a more significant increase in spouses' total working hours for spouses with no chronic disease (significant 125 hours/year) than those with at least one chronic disease (insignificant 57 hours per year). We find the same trend in their hourly income as well (4.0 significant compared with 0.7 insignificant), indicating that health is a critical variable for promoting working productivity.

[Insert Table 10 here]

[Insert Table 11 here]

Mental health status is also critical in individuals' labor provision decisions. We examine whether the effects of LTCI on spouses' labor supply become smaller or insignificant after adding the variable of mental health level (measured by the depression score of the CES-D scale) to the covariates. Table 12 confirms this mediating mechanism, from which we find a smaller or a decreased significant level after we controlled for spouses' mental health level. In the previous literature, people with a score less than 7.5 generally have no depressive symptoms; people with a score between 7.5 and 10 may have depressive symptoms; people with a score of 10 or more have depressive symptoms and it is recommended to see a psychologist (Radloff, 1977). In Table 13, we find a more significant increase in spouses' probability of labor provision, total yearly working hours and hourly wages with a higher mental health level (9.7% significant compared with 3.1% insignificant; 142 hours/year significant compared with 34 hours/year insignificant; 6.7 yuan/hour significant compared with -0.03 yuan/hour insignificant).

The results of T-test suggest that physical and mental health of caregivers play important roles in the significant promotion of spouses' labor market performance after the introduction of LTCI. Through a positive spillover effect of LTCI on spouses' health status, their labor supply and productivity would be further stimulated.

[Insert Table 12 here]

[Insert Table 13 here]

B.2 Mechanism II: Spillover Effects Through the Re-allocation of Time

Many researchers argue that the complementary relationship between leisure and care is essential when modeling spouses' labor supply choices (McGarry, 2006; Michaud and Vermeulen, 2011; Jeon and Pohl, 2017). Rational individuals with 24 hours per day face the choice of allocating time to the following three categories of activities: labor supply (L), enjoyment of leisure (including sleeping, eating, socializing, and entertainment, E), and provision of informal care (C). These three items are strictly mutually exclusive as one cannot enjoy leisure or provide informal care services while working. However, there is an unintuitive relationship between care activities and leisure time. McGarry (2006) finds that female caregivers often cut back on leisure time in order to provide care, while Michaud and Vermeulen (2011) argue that individuals will enjoy more leisure time to accompany their disabled spouses. Jeon and Pohl (2017) find a tendency to spend more leisure time together for couples after a severe health shock, reflecting a complementarity relationship of leisure and informal care. From this perspective, LTCI may reduce one's leisure time at the same time while reducing their informal care provisions. It can be seen from Table 14 that the implementation of LTCI has a negative effect on spouses' total hours of providing informal care to their disabled family members, and their leisure hours are also reduced. We find evidence that LTCI crowds out about 81.62 hours of informal care per year¹⁷, and the yearly leisure time is also taken away to some extent (a 238.46 hours' decrease in yearly leisure hours). LTCI provides an opportunity for spouses of the disabled to get away from the long-term care burden and obtain more time at their disposal. Moreover, complementary spousal leisure predicts a decrease in spouses' leisure time and a further increase in their labor supply (Boyle and Lahey, 2016).

[Insert Table 14 here]

¹⁷ Descriptive statistics shows that the average informal care time for the treatment group is 117.043 hours/year. After the LTCI's introduction, the length of informal care was reduced by 81.62 hours. It could be found through simple calculations that informal care time was reduced from 137.448 hours per year [calculated from $(117.043 \times 4 + 81.62) / 4$] to 55.828 hours per year [calculated from $(137.448 - 81.62)$] with a reduction rate of 59.4% [calculated from $(81.62 / 137.448 \times 100\%)$], indicating that LTCI greatly reduced caregivers' burden of informal family care.

From the above discussion, we find that after the introduction of LTCI, spouses' labor supply increase could be partly attributed to the reduction of both informal care time and leisure time. Individuals reported needing an average of 7.1 hours of sleep to feel their best and people with less than 7 hours of sleep time per day are often faced with time conflicts (Knutson et al., 2017). We examine whether the effects of LTCI on spouses' labor supply become smaller or insignificant after adding the variable of the leisure time (measured by the yearly leisure hours) to the covariates. Table 15 confirms this mediating mechanism, from which we find a smaller or a decreased significant level after we controlled for spouses' leisure time. Moreover, we divided the respondents into two groups according to the leisure time (including sleeping, eating, socializing, and entertainment) less than 8 hours and at least 8 hours (Rosenzweig, 1985). In Table 16, we find that compared with caregivers who prefer and spend more leisure time, spouses who were faced with time conflicts before the introduction of LTCI tend to re-enter the labor market (10.7%) once they have access to more disposable daily hours. However, individuals who choose to spend more leisure time with their families are less likely to face a severe time conflict between family care provision and making money in labor market. As a result, we find a smaller increase of the labor participation rate (6.1%).

[Insert Table 15 here]

[Insert Table 16 here]

VI. Heterogeneous Analysis

A. Heterogeneity of Individual's Characteristics

1) Age

In most of the literature, labor supply is measured among caregivers aged 15-64 years old in the light of corresponding mandatory retirement legislation. In China, the current retirement age is 60 years old for men, 55 years old for female officials, and 50 years old for female workers (State Council of China, 1994). However, under the current demographic changes, excluding people older than 65 from analyses overlooks an increasing extent to which the elderly enters the labor force. The latest statistics in Japan and the US reveal labor participation rates at 42.7% and 30.8% for people aged 65-69 years old to participate in the labor force market (Statistics, 2016; Japan, 2016). It can be reasonably expected that the labor participation rate of the elderly in a gradually aging society will increase in the future. More importantly, as many caregivers in China are spouses of disabled elderly individuals (Insurance Association of China, 2019), analysis of labor supply among caregivers aged 65+ will help inform family and labor policy formulations in the context of a super-aged society.

In this section, we first conduct the propensity score matching of different individuals at their age of 59 years old and below (younger people) and 60 years old and above (older people). We expect a different spillover effect in different age groups at their distinctive stages of life and have different attachment levels to the labor market (Fu et al., 2017) and their distinctive elasticity of labor supply would impact the strength of LTCI's spillover effects. Notable differences of confounding covariates between treatment and control groups are commonly found across pre-PSM subsamples (Table C.2.1 and Table C.2.2). All the covariates are balanced in post-PSM subsamples (i.e., mean SD < 10%).

The heterogeneity of LTCI's spillover effect by age is shown in Table 17. PSM-DID estimators find that LTCI does improve the labor participation rate of treatment group for different age groups with 6.3% in the younger group (59 years old and below) and 5.2% in the older group (60+ years old), indicating that young spouses are a bit more sensitive to the benefits of LTCI. However, treated older individuals contribute to the majority of the prolonged working hours (194 hours/year) while we do not find a significant difference in younger adults' labor hours. In

addition, things are different with regard to subdivided types of work. After the introduction of LTCI, treated younger people become more devoted to self-employed work (with an increase of 10.5% in the labor participation rate and a prolonged 122 hours' labor hours per year). Meanwhile, treated older people tend to increase their paid employment hours (an increase of 99.7 hours/year). Results are interesting in that we find different strategies to promote their labor supply in different age groups. For the relatively younger (probably healthier) people, reducing self-employed labor hours to provide informal family care is a rather easy and convenient way to take care of patients. Accordingly, we find a significant increase in their self-employed working hours once LTCI shared the caring pressure with them. However, younger adults in urban areas are more likely to pursue a promising future career in organizations in their later lives. For them, devoting their extra-saved time by LTCI to improve their working productivity or pursue a job with a higher salary would be a wiser choice (Bauer and Sousa-Poza, 2015). In contrast, for individuals who are close to or exceed the statutory retirement age, we find an increase only in the paid employment working hours. As they grow older, individuals might find themselves easier to get tired in the provision of self-employed work (especially for self-employed agricultural work), and some of them increase their paid employment work as long as they got the opportunity to be re-employed or re-enter the labor market, resulting in an increase in paid employment working hours.

[Insert Table 17 here]

2) Gender

Previous studies show significant heterogeneity between groups of different genders as regards the spillover effects of LTCI. For example, women are more likely to become caregivers than men, and LTCI may thus have a more significant impact on wives' labor supply than husbands'. Moreover, female caregivers are more likely to provide intensive care for their parents, resulting in a lower probability of labor participation than non-caregivers (Skira, 2015). Compared to older and female people, young and male usually have a higher probability to work (Twenge, 2010). However, an increasing number of Chinese women are highly educated and employed as full-time workers in the last decades. Therefore, we expect that the spillover effects will be larger when the time conflict binds, especially for highly educated female caregivers. In this section, we investigate the difference in spillover effects of LTCI between different gender groups. We first conduct two propensity score matchings for male and female individuals and obtain each covariate's balancing test. Notable differences of confounding covariates between treatment and control groups are commonly found across pre-PSM subsamples, but all the covariates are balanced in post-PSM subsamples (mean SD < 10%). The spillover effects of LTCI on wives' and husbands' labor supply are listed in Table 13.

From Table 18, we find that LTCI does improve the labor participation rate of the treatment group, particularly for wives. LTCI treated females are 8.9% more likely to participate in the labor market than untreated women, while that number of the male group is 2.4% (insignificant). However, things are different concerning subdivided types of work. After the introduction of LTCI, treated males become more devoted to self-employed work. They become 7.5% more likely to re-enter the self-employed labor market or prolong their self-employed work by 120.6 hours/year (compared with 6% and 71.4 hours/year for female). However, treated females tend to increase their paid hours employed by government, organizations, and firms (with an increase of labor participation rate of 4.5% and a prolonged 107.9 hours/year in paid employment). The last decades have seen an increasing number of Chinese women being highly educated and employed as full-time workers. For them, balancing working and caregiving is more stressful than that for traditional housewives (Makita, 2010). The success of LTCI in encouraging wives to re-enter/remain in the labor force is a good lesson for other policies aiming to stimulate women's labor supply. Besides, as more spare time returns to the caregivers, they will have more continuous-time, and we expect that their probability of finding a job with a fixed schedule increases.

[Insert Table 18 here]

B. Heterogeneity of LTCI Policy

1) Heterogeneity of LTCI's target people

Previous studies have found that formal care is not necessarily a substitute for informal family care (Fischer and Müller, 2020). As the degree of disability of the elderly continues to deepen, the substitute relationship between family care and formal care is gradually weakening, and it shows a complementary relationship in the special end-stage before the death of the elderly (Huang and Fu, 2017; Bonsang, 2009). Therefore, the impact of LTCI on the labor supply of caregivers of moderately and severely disabled people is expected to be different. Most of the LTCI pilots in China cover severely disabled persons and long-term disabled persons only. However, other cities also cover moderately disabled persons and demented persons. In order to characterize this policy feature, we introduce a dummy variable $policy_i^a$ with $policy_i^a = 0$ for pilots covering severely disabled people only (include individuals with a Barthel score ≤ 40 only before July 2018) and $policy_i^a = 1$ for pilots covering both moderately and severely disabled people (include individuals with a Barthel score ≤ 40 or >40 before July 2018). We demonstrate the result of PSM-DID method in two groups of LTCI pilot cities and find a more significant increase in spouses' labor supply in cities with both moderately and severely disabled covered than that of cities with eligibilities only for severely disabled.

[Insert Table 19 here]

Table 19 shows that the heterogeneity in LTCI eligibility has a significantly positive effect on spouses' labor participation (6.8% compared with 4.6%) and yearly working hours (149.2 hours/year compared with an insignificant 43.4 hours/year), indicating that in pilot cities that only cover severely disabled individuals, the labor supply of spouses has not been completely released since individuals with severe disabilities usually require both formal care and informal care at the same time. For LTCI pilot cities that cover both the severely and moderately disabled, LTCI has a more obvious spillover effect on individuals' labor supply since moderately disabled people only need one of the formal care and family informal care at a time, making informal care simply a substitute for formal care.

2) Heterogeneity of LTCI's benefit rules

In China, some LTCI pilots pay benefits in kind, which reimburse only a certain percentage of expense spent on institutional formal care, while other pilots pay benefits in cash at the same time. Individuals are usually allowed to choose one of the two forms of benefits above and caregivers of the disabled would make their care and career choice accordingly. The acceptance of benefits in cash usually indicates that the disabled are not permitted to the cheaper formal care services. Therefore, caregivers in LTCI pilot cities providing cash benefits will not necessarily reduce their informal care hours. We cannot expect the labor supply to be affected significantly by LTCI in these cities. For pilot cities that do not provide cash benefits, disabled individuals can enjoy LTCI benefits as long as they purchase the exact formal care services. As a result, benefits in kind may crowd out more informal care and promote family caregivers' labor supply to a larger extent. In order to characterize this policy feature, we introduce a dummy variable $policy_{it}^b$ with $policy_{it}^b = 1$ for pilots providing benefit in cash while $policy_{it}^b = 0$ for pilots who don't. We demonstrate the result of PSM-DID method in two groups of LTCI pilot cities and find a significant decrease in spouses' labor supply in cities with both benefit in kind and in cash provided.

[Insert Table 20 here]

Table 20 shows the results of the difference among individuals in the treatment and control groups before and after the introduction of differently designed LTCI policies. We find that compared with pilot cities that only provide benefits in kind, there is a smaller or an insignificant increase in spouses' labor supply in LTCI pilot cities that provide benefit in cash. Spouses may continue to provide the same amount of informal care hours as before while receiving subsidies in cash from LTCI at the same time. Therefore, in pilot cities that provide benefits in cash, the labor supply of spouses has not been completely released. In contrast, spouses in LTCI pilot cities with only benefits in service provided demonstrate larger and more significant spillover effects of LTCI (an increase of 7.7% in labor participation rate and a prolonged 140.3 hours/year). We find the same phenomenon in the sub-categories of work (self-employed work / employed work). However, we find a significant decrease in the self-employed non-agricultural work for individuals in LTCI pilots offering cash benefits (not shown in Table 20). For individuals with a severe disability, LTCI subsidies in cash act as a stable and continuous cash flow for the disabled adult's whole family. Spouses get a chance to breathe between the busy care activities and plan for future career. They may reduce their self-employed working hours and seek an opportunity for full-time paid employ work.

VII. Robustness Checks

In the above sections, we identify the spillover effects of LTCI on spouses' labor supply by conducting the PSM-DID method. We regard the treatment group and the control group randomly assigned and have the same trend in the outcome variables and the mediation variables. However, there are other problems that need to be clarified to ensure the credibility of our results. Firstly, the spillover effects of LTCI on spouses' labor supply may differ across 42 pilot cities. We try to figure out whether there are significant differences in the spillover effects of LTCI under different socio-economic conditions among different cities. Therefore, we conduct 42 separate PSM-DIDs, selecting individuals in one LTCI pilot city as the treatment group at one time. The control group is still cities that have not implemented LTCI before July 2018. Secondly, contingency may occur in LTCI pilot cities. An accidental behavior of an individual in the treatment group may cause great fluctuations in the outcome variables. Therefore, we censored the outcome variables at the 95% level or take the

logarithm before conducting PSM-DID. Thirdly, we conduct a placebo test (or falsification test) in years with no LTCI launched (e.g., in 2013 and 2015). Fourthly, the LTCI pilots are often cities with a higher socio-economic condition, a relatively sufficient basic medical insurance fund, and a more severe aging problem. Here, we replace the control group with individuals in cities where LTCI was launched between August 2018 and April 2021 (34 cities). The treatment group is still individuals in cities that implemented LTCI before July 2018. Therefore, the city-level characteristics of the treatment group and control group become more balanced. Fifthly, the increase in individuals' labor supply may be caused by other reforms in basic medical insurance and have nothing to do with the introduction of LTCI. Therefore, we explore the difference in the frequency of basic medical insurance use between treatment and control groups to rule out the probability that other potential factors instead of LTCI induced the increased labor supply. Finally, previous studies use different matching functions and approaches (Jeon and Pohl, 2017). Here we will conduct “K-Nearest Neighbor Matching” and “Coarsened Exact Matching (CEM)” Difference-in-Difference and get the results under different matching hypotheses to verify the conclusions we drew before.

A. Considerations of the Differences in Spillover Effects Across All LTCI Pilots

The spillover effects of LTCI on spouses' labor supply may differ across pilot cities. In this section, we conduct 13 separate PSM-DIDs within 13 provinces¹⁸. We select individuals in one LTCI pilot city as the treatment group at a time and individuals in the rest cities of the province as the new control group. Table 21 and Fig. 7 show the treatment groups and control groups in the 13 PSM-DIDs. Fig. 8 depicts the spillover effects of LTCI within 13 provinces. It can be found that the spillover effects in each treatment group fluctuate in a small range around the spillover effects in the overall sample in our fundamental analysis. Fig. 8 verifies that our conclusion is generally held in most of the LTCI pilot cities.¹⁹

¹⁸ In order to avoid the impact of accidental factors, we rule out the pilot city with a sample size of less than 50 in the treatment group.

¹⁹ Theoretically, the small sample size of the treatment group will not result in a significant bias in our PSM-DID results. There is no additional requirement on sample size as long as covariates pass the balancing test of matching. In fact, in the classic research of Moser and Voena (2012), the sample size of the treatment group was only 4.6%.

[Insert Table 21 here]

[Insert Figure 7 here]

[Insert Figure 8 here]

B. Considerations of the Contingency That May Occur on LTCI Pilots

Accidental behaviors of an individual in the treatment group may cause great fluctuations in the outcome variables. Therefore, contingencies that occurred in LTCI pilot cities will make the results deviate from their actual value. We censored the outcome variables at the 98% level (delete extreme values and only reserve samples whose outcome variables fall between its [1%, 99%] interval) or take the logarithm of them and repeat the procedure of PSM-DID. Results are demonstrated in Table 22 and Table 23. We find that the main conclusions we draw in the fundamental analysis still hold after the censoring and taking logarithm of outcome variables.

C. Placebo Tests on Other Years

Since almost all of the LTCI pilots were launched during 2016 to 2018 (LTCI was first introduced in July 2012 for Qingdao, January 2015 for Weifang, May 2015 for Changchun), we rule out individuals in Qingdao city, Weifang city, and Changchun city and select 2012 and 2014 as the new hypothetical year for the introduction of LTCI as a placebo test. If the increase in labor supply is caused by the introduction of LTCI, we could reasonably expect that the results of placebo tests will not be significant. From Table 24, we find that the results of the outcome variables are all insignificant, indicating that they all passed the placebo test in 2012 and 2014. Therefore, the increase in labor participation rate, yearly working hours, and hourly income in our fundamental analysis should be reasonably attributed to the introduction of LTCI.

[Insert Figure 9 here]

D. Considerations of the Different Socio-economic Conditions Between Treatment and Control Groups

The LTCI pilots are often cities with a higher socio-economic condition, a relatively sufficient basic medical insurance fund, and a more severe aging problem. Here, we replace the control group with individuals in cities where LTCI was launched between August 2018 and April 2021 (34 cities²⁰). The treatment group is still individuals in cities that implemented LTCI before July 2018 (42 cities). Therefore, the city-level characteristics of the treatment group and control group become more balanced. Table 25 shows the spillover effects of LTCI in pilot cities with LTCI launched before July 2018, or after July 2018 and before April 2021. We find a similar spillover effect between these selected cities, verifying our results' robustness in the fundamental analysis.

[Insert Table 22 here]

[Insert Table 23 here]

[Insert Table 24 here]

[Insert Table 25 here]

²⁰ These cities are: Hezhou, Dezhou, Yangzhou, Zaozhuang, Qinhuangdao, Urumqi, Wenzhou, Meihokou, Hunchun, Tianmen, Yichang, Dingzhou, Gannan Prefecture, Qianxinan Prefecture, Xiangtan, Tianjin, Fuzhou, Kunming, Hanzhong, Jincheng, Kaifeng, Panjin, Nanning, Yiwu, Changzhou, Wuxi, Zhoushan, Manzhouli, Hulun Buir, Taizhou (in Jiangsu), Tangshan, Shijiazhuang, Wuhai, and Jinjiang.

E. Considerations of the Probably Changed Habits of Basic Medical Insurance Usage after the Introduction of LTCI

The increase in individuals' labor supply may be caused by other changes in basic medical insurance and have nothing to do with the introduction of LTCI. Therefore, we explore the difference in the frequency of basic medical insurance use between treatment and control groups to rule out the probability that the increased labor supply was induced by other potential factors instead of LTCI. Changes in the frequency of basic medical insurance use before and after the introduction of LTCI are shown in Table 26. We do not find a significant increase or decrease in individuals' medical habits, indicating that the increase in spouses' labor supply should be reasonably attributed to LTCI.

F. Spillover Effects of LTCI under Other Matching Framework

Previous studies use different matching functions and approaches before conducting a DID framework. Matching is a non-parametric method to control the difference between the treatment and control groups by screening samples. Here we conduct a “K-Nearest Neighbor Matching (K-NNM)” and a “Coarsened Exact Matching (CEM)” Difference-in-Difference framework and get the results under the corresponding matching hypotheses to verify the conclusions we drew before.

Firstly, to balance covariates between the treatment and control groups, we conduct a K-Nearest Neighbor Matching (K-NNM) (Abadie et al., 2004). NNM uses nearest neighbor matching across the covariates to estimate the average treatment effect on outcome variables by comparing outcomes between treated and control observations. This program pairs observations to the closest K matches in the opposite treatment group to provide an estimate of the counterfactual treatment outcome. It allows for matching over a multi-dimensional set of variables, giving options for the weighting matrix to be used in determining the optimal matches. It also allows exact matching on a subset of variables. In addition, it allows for bias correction of the treatment effect, and estimation of either the sample or population variance, with or without assuming a constant treatment effect. Finally, it allows observations to be used as a match more than once, thus making the order of matching irrelevant.

Secondly, we conduct a Coarsened Exact Matching (CEM) Difference-in-Difference framework (Iacus et al., 2012; Blackwell et al., 2009). CEM

is a method for improving the estimation of causal effects by reducing imbalance in covariates between treated and control groups. Coarsened exact matching requires fewer assumptions and possesses more attractive statistical properties than many other matching methods. In particular, CEM estimates of causal effects are less sensitive to model specification choice (Stuart, 2010). CEM bounds the degree of model dependence²¹ and causal effect estimation error by ex-ante user choice (Iacus et al., 2012), and reducing the maximum imbalance on one variable has no effect on others. Moreover, CEM meets the congruence principle and does not require a separate procedure to restrict data to common support. It is approximately invariant to measurement error, and balances all nonlinearities and interactions in sample (i.e., not merely in expectation). Table 27 shows the spillover effects of LTCI under a “K-Nearest Neighbor Matching” framework (K=4) and a “Coarsened Exact Matching (CEM)” framework. We find that the spillover effects of LTCI remain the same under different assumptions and various matching methods, verifying the robustness of our fundamental analysis.

[Insert Table 26 here]

[Insert Table 27 here]

²¹ Model dependence is defined by how much the predicted value of the outcome variable varies as a function of the statistical model for a given set of explanatory variables (Stuart, 2010).

VIII. Conclusions

Aging problems pose serious challenges for the health and long-term care system in many countries around the world. In this research, we show the significant and positive spillover effects of the LTCI on caregivers' labor supply both on the extensive and intensive margin. The family labor force is liberated after the introduction of LTCI, and poverty alleviation could be promoted through an increase in the employment rate. A promotion of 5.9% is found on spouses' labor participation rate, indicating a significant reduction in care burden on Chinese caregivers. Moreover, we found that the increase in working hours is different among various types of work. We see increases in individuals' total hours of work (80.78 hours/year), the hours of self-employed work (96.23 hours/year), and the hours of paid employment (67.28 hours/year). We expect a further increase in people's paid employment as the adjustment of these jobs' time schedule needs a longer time to complete. Moreover, hourly income was also promoted in paid employment (2.76 RMB yuan/hour), which could be explained by the improvement of caregivers' productivity and their pursuit of a job promotion (Carmichael and Charles, 2003). The success of Chinese LTCI on stimulating individuals' labor supply is intuitive in the super-aged society, regarding the increase in caregivers for frail old persons, which provides an excellent example to other countries where encouraging caregivers' labor supply is a priority.

We also investigate the mediation mechanism between LTCI and caregivers' labor supply. LTCI covers the daily care expenses for the disabled elderly caused by illness or disability, which is expected to reduce the physical toil and psychological pressure of the caregivers. We find that LTCI in China does improve individuals' health level measured by the number of chronic diseases and the depression score. After liberated from heavy care duty, caregivers become more satisfied and optimistic toward lives, and their risk of depression would be reduced (Dáz et al., 2020). Meanwhile, caregivers' physical discomfort and their occurrence of chronic diseases are also mitigated by reducing physical toil from care burden, resulting in a higher enthusiasm and productivity in the workplace. Moreover, from the perspective of time re-allocating, individuals' informal care hours, as well as their leisure time, are reduced at the same time after the introduction of LTCI. By cutting down the time spent on odd and trivial matters (e.g., caring hours and social entertainment), individuals obtain more continuous and undisturbed time to seek career development,

which increases caregivers' working hours (and productivity at the same time).

Age and gender differences are found regarding the spillover effects of LTCI on labor supply. We find that younger individuals are more sensitive to the subsidization of formal caregiving by LTCI in that LTCI improves the labor participation rate of treatment group with 6.3% in younger people group (59 years old and below) and 5.2% in older people group (60 years old and above). Moreover, young people tend to become more devoted to the self-employed work (an increase of 10.5% in the probability of participating self-employed work and 122 hours/year in the prolonged self-employed working hours). Meanwhile, treated older people tend to increase their paid employment hours (99.7 hours/year). Previous literature reveals that younger usually devote their spare time improve their working production (Bauer and Sousa-Poza, 2015), while older ones do not expect an extra promotion in their career prospects and find themselves more suitable to prolong their working hours. Moreover, results are interesting in that we find different strategies to promote their labor supply in different gender groups. LTCI increases the labor participation rate both on males and females, while the promotion on women's labor participation rate (8.9%, significant) is larger than that of men (2.4%, insignificant), mainly because most of the care responsibilities were borne by women before the introduction of LTCI. In addition, spouses usually adopt different strategies to increase their labor hours. We see a larger increase in men's probability of participating in self-employed work (7.5%) and the corresponding working hours (120.6 hours/year) than female (6.0% and 71.4 hours/year), indicating that policies aiming to stimulate labor supply should be formulated separately for men and women. By contrast, women's probability of participating in paid employment (4.5%) and their hours (107.9 hours/year) increased compared to females in non-LTCI cities. For policymakers, LTCI's heterogeneous spillover effects on different age and gender groups should be taken into consideration.

Considering differences in LTCI policy design regarding target group and benefit rules, we study the heterogeneous spillover effects among different pilot cities. We find evidence that compared with severely disabled elderly, informal care for moderately disabled elderly demonstrates a stronger substitution relationship with formal care, making the spillover effects greater among cities targeting both moderately and severely disabled than those only targeting severely ones. Besides, we find smaller spillover effects of LTCI on labor supply in pilots providing benefit in

cash than those who do not, mainly because that cash allowance only attributes to caregivers' non-wage income, but has little impact on the opportunity cost of providing informal care by family members (Geyer and Korfhage, 2015a; Carmichael and Charles, 2003). Evidence shows that the LTCI pilots with benefits in kind would stimulate caregivers' labor supply much more than pilots with benefits in cash.

However, this research has some limitations. The first limitation is a concern regarding the kernel PSM method on addressing endogeneity. Since the matching accounts for observable covariates, unobservable influences still may remain in the model. Other researchers try to tackle the issue with the IV method using family characteristics (such as the number of siblings or parental health) as instruments (Heitmueller, 2007; Van Houtven et al., 2013). Further analyses with Chinese data using the IV method are required for a cohesive picture of the spillover effects of LTCI. The second shortcoming lies in that considering the short implementation time of LTCI, the effects of policies cannot be fully demonstrated at present. Further studies could use updated panel data to detail the lagging effects of LTCI policies.

References

- [1] Abadie, A., Herr, J. L., Imbens, G., & Drukker, D. M. (2004). *NNMATCH: Stata module to compute nearest-neighbor bias-corrected estimators*.
- [2] Ågren, S., Evangelista, L., & Strömberg, A. (2010). Do partners of patients with chronic heart failure experience caregiver burden? *European Journal of Cardiovascular Nursing*, 9(4), 254–262.
- [3] Aguiar, M., Bils, M., Charles, K. K., & Hurst, E. (2021). Leisure luxuries and the labor supply of young men. *Journal of Political Economy*, 129(2), 000–000.
- [4] Alfakhri, A. S., Alshudukhi, A. W., Alqahtani, A. A., Alhumaid, A. M., Alhathlol, O. A., Almojali, A. I., Alotaibi, M. A., & Alaqeel, M. K. (2018). Depression among caregivers of patients with dementia. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 55, 0046958017750432.
- [5] Andersen, M. (2015). Heterogeneity and the effect of mental health parity mandates on the labor market. *Journal of Health Economics*, 43, 74–84.
- [6] Arai, Y., & Zarit, S. H. (2011). Exploring strategies to alleviate caregiver burden: Effects of the National Long-Term Care insurance scheme in Japan. *Psychogeriatrics*, 11(3), 183–189.
- [7] Ariizumi, H. (2008). Effect of public long-term care insurance on consumption, medical care demand, and welfare. *Journal of Health Economics*, 27(6), 1423–1435.
- [8] Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399–424.
- [9] Bakx, P., De Meijer, C., Schut, F., & Van Doorslaer, E. (2015). Going formal or informal, who cares? The influence of public long-term care insurance. *Health Economics*, 24(6), 631–643.
- [10] Barnay, T., & Juin, S. (2016). Does home care for dependent elderly people improve their mental health? *Journal of Health Economics*, 45, 149–160.
- [11] Barón, J. D., & Cobb-Clark, D. A. (2010). Occupational segregation and the gender wage gap in private-and public-sector employment: A distributional analysis. *Economic Record*, 86(273), 227–246.
- [12] Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical

- considerations. *Journal of Personality and Social Psychology*, 51(6), 1173.
- [13] Bauer, J. M., & Sousa-Poza, A. (2015). Impacts of informal caregiving on caregiver employment, health, and family. *Journal of Population Ageing*, 8(3), 113–145.
- [14] Bhattacharyya, O., Delu, Y., Wong, S. T., & Bowen, C. (2011). Evolution of primary care in China 1997–2009. *Health Policy*, 100(2–3), 174–180.
- [15] Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). cem: Coarsened exact matching in Stata. *The Stata Journal*, 9(4), 524–546.
- [16] Blaise, M. & others. (2018). *The informal care provision: What are the genuine incentives of children?* Bureau d’Economie Théorique et Appliquée, UDS, Strasbourg.
- [17] Blundell, R., Costa Dias, M., Meghir, C., & Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5), 1705–1753.
- [18] Blundell, R., MaCurdy, T., & Meghir, C. (2007). Labor supply models: Unobserved heterogeneity, nonparticipation and dynamics. *Handbook of Econometrics*, 6, 4667–4775.
- [19] Blundell, R., Pistaferri, L., & Saporta-Eksten, I. (2016). Consumption inequality and family labor supply. *American Economic Review*, 106(2), 387–435.
- [20] Bobinac, A., Van Exel, N. J. A., Rutten, F. F., & Brouwer, W. B. (2010). Caring for and caring about: Disentangling the caregiver effect and the family effect. *Journal of Health Economics*, 29(4), 549–556.
- [21] Bolin, K., Lindgren, B., & Lundborg, P. (2008). Informal and formal care among single-living elderly in Europe. *Health Economics*, 17(3), 393–409.
- [22] Bonsang, E. (2009). Does informal care from children to their elderly parents substitute for formal care in Europe? *Journal of Health Economics*, 28(1), 143–154.
- [23] Boyle, M. A., & Lahey, J. N. (2016). Spousal labor market effects from government health insurance: Evidence from a veterans affairs expansion. *Journal of Health Economics*, 45, 63–76.
- [24] Bracke, P., Christiaens, W., & Wauterickx, N. (2008). The pivotal role of women in informal care. *Journal of Family Issues*, 29(10), 1348–1378.
- [25] Bredtmann, J., & Vonnahme, C. (2019). Less money after divorce—how the 2008 alimony reform in Germany affected spouses’ labor supply, leisure and marital stability. *Review of Economics of the Household*, 17(4), 1191–1223.

- [26] Breen, R., Karlson, K. B., & Holm, A. (2013). Total, direct, and indirect effects in logit and probit models. *Sociological Methods & Research*, 42(2), 164–191.
- [27] Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, 163(12), 1149–1156.
- [28] Bryson, A., Dorsett, R., & Purdon, S. (2002). *The use of propensity score matching in the evaluation of active labour market policies*.
- [29] Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72.
- [30] Carmichael, F., & Charles, S. (1998). The labour market costs of community care. *Journal of Health Economics*, 17(6), 747–765.
- [31] Carmichael, F., & Charles, S. (2003). The opportunity costs of informal care: Does gender matter? *Journal of Health Economics*, 22(5), 781–803.
- [32] Casado-Marín, D., García-Gómez, P., & López-Nicolás, Á. (2011). Informal care and labour force participation among middle-aged women in Spain. *SERIEs*, 2(1), 1–29.
- [33] Chang, S., Yang, W., & Deguchi, H. (2020). Care providers, access to care, and the Long-term Care Nursing Insurance in China: An agent-based simulation. *Social Science & Medicine*, 244, 112667.
- [34] Charles, K. K., & Sevak, P. (2005). Can family caregiving substitute for nursing home care? *Journal of Health Economics*, 24(6), 1174–1190.
- [35] Chatterji, P., Alegria, M., Lu, M., & Takeuchi, D. (2007). Psychiatric disorders and labor market outcomes: Evidence from the National Latino and Asian American Study. *Health Economics*, 16(10), 1069–1090.
- [36] Chen, H., Qian, Y., Dong, Y., Yang, Z., Guo, L., Liu, J., Shen, Q., & Wang, L. (2020). Patterns and changes in life expectancy in China, 1990–2016. *PloS One*, 15(4), e0231007.
- [37] Chiang, W., & Wu, M. (2020). A study on the development and transformation of caregivers' work values. *International Joint Conference on Arts and Humanities (IJCAH 2020)*, 1449–1452.
- [38] Chu, K., Liu, D., & Duan, W. (2018). The mechanism of rural households' smooth consumption under health shocks—Also on the relationship between external security and household self-protection (in Chinese). *Nankai Economic*

Research, 200(02), 39–55.

- [39] Ciani, E. (2012). Informal adult care and caregivers' employment in Europe. *Labour Economics*, 19(2), 155–164.
- [40] Coe, N. B., & Van Houtven, C. H. (2009). Caring for mom and neglecting yourself? The health effects of caring for an elderly parent. *Health Economics*, 18(9), 991–1010.
- [41] Cremer, H., Lozachmeur, J.-M., & Pestieau, P. (2016). The design of long term care insurance contracts. *Journal of Health Economics*, 50, 330–339.
- [42] Crespo, L., & Mira, P. (2014). Caregiving to elderly parents and employment status of European mature women. *Review of Economics and Statistics*, 96(4), 693–709.
- [43] Davidoff, T. (2010). Home equity commitment and long-term care insurance demand. *Journal of Public Economics*, 94(1–2), 44–49.
- [44] Dentinger, E., & Clarkberg, M. (2002). Informal caregiving and retirement timing among men and women: Gender and caregiving relationships in late midlife. *Journal of Family Issues*, 23(7), 857–879.
- [45] Díaz, A., Ponsoda, J. M., & Beleña, A. (2020). Optimism as a key to improving mental health in family caregivers of people living with Alzheimer's disease. *Aging & Mental Health*, 24(10), 1662–1670.
- [46] Dispenza, F., Brennaman, C., Harper, L. S., Harrigan, M. A., Chastain, T. E., & Procter, J. E. (2019). Career development of sexual and gender minority persons living with disabilities. *The Counseling Psychologist*, 47(1), 98–128.
- [47] Do, Y. K., Norton, E. C., Stearns, S. C., & Van Houtven, C. H. (2015). Informal care and caregiver's health. *Health Economics*, 24(2), 224–237.
- [48] Dolls, M., Doerrenberg, P., Peichl, A., & Stichnoth, H. (2018). Do Retirement Savings Increase in Response to Information About Retirement and Expected Pensions? *Journal of Public Economics*, 158, 168–179. <https://doi.org/10.1016/j.jpubeco.2017.12.014>
- [49] Ettner, S. L. (1995). The impact of “parent care” on female labor supply decisions. *Demography*, 32(1), 63–80.
- [50] Ettner, S. L. (1996). New evidence on the relationship between income and health. *Journal of Health Economics*, 15(1), 67–85.
- [51] Feng, J., Wang, Z., & Yu, Y. (2020). Does long-term care insurance reduce hospital utilization and medical expenditures? Evidence from China. *Social*

Science & Medicine, 258, 113081.

- [52]Feng, Z., Glinskaya, E., Chen, H., Gong, S., Qiu, Y., Xu, J., & Yip, W. (2020). Long-term care system for older adults in China: Policy landscape, challenges, and future prospects. *The Lancet*, 396(10259), 1362–1372. [https://doi.org/10.1016/S0140-6736\(20\)32136-X](https://doi.org/10.1016/S0140-6736(20)32136-X)
- [53]Fischer, B., & Müller, K.-U. (2020). Time to care? The effects of retirement on informal care provision. *Journal of Health Economics*, 73, 102350.
- [54]Fisher, L. D., Dixon, D. O., Herson, J., Frankowski, R. F., & Pearce, K. (1990). Analysis of randomized clinical trials: Intention to treat. *Statistical Issues in Drug Research and Development*, 331, 344.
- [55]Fu, R., Noguchi, H., Kawamura, A., Takahashi, H., & Tamiya, N. (2017). Spillover effect of Japanese long-term care insurance as an employment promotion policy for family caregivers. *Journal of Health Economics*, 56, 103–112. <https://doi.org/10.1016/j.jhealeco.2017.09.011>
- [56]Fukahori, R., Sakai, T., & Sato, K. (2015). The effects of incidence of care needs in households on employment, subjective health, and life satisfaction among middle-aged family members. *Scottish Journal of Political Economy*, 62(5), 518–545.
- [57]Ganapathy, V., Graham, G. D., DiBonaventura, M. D., Gillard, P. J., Goren, A., & Zorowitz, R. D. (2015). Caregiver burden, productivity loss, and indirect costs associated with caring for patients with poststroke spasticity. *Clinical Interventions in Aging*, 10, 1793.
- [58]Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, 49(5), 1701–1720.
- [59]Garthwaite, C., Gross, T., & Notowidigdo, M. J. (2014). Public health insurance, labor supply, and employment lock. *The Quarterly Journal of Economics*, 129(2), 653–696.
- [60]Geyer, J., Haan, P., & Korfhage, T. (2017). Indirect fiscal effects of long-term care insurance. *Fiscal Studies*, 38(3), 393–415.
- [61]Geyer, J., & Korfhage, T. (2015a). Long-term care insurance and carers' labor Supply—A structural model. *Health Economics*, 24(9), 1178–1191.
- [62]Geyer, J., & Korfhage, T. (2015b). Long-term care reform and the labor supply of household members—evidence from a quasi-experiment. *Available at SSRN*

2706538.

- [63]Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, 2(3), 109.
- [64]Hajek, A., Lehnert, T., Wegener, A., Riedel-Heller, S. G., & König, H.-H. (2018). do individuals in old age prepare for the risk of long-term care? Results of a population-based survey in Germany. *International Journal of Environmental Research and Public Health*, 15(10), 2189.
- [65]Hallberg, D. (2003). Synchronous leisure, jointness and household labor supply. *Labour Economics*, 10(2), 185–203.
- [66]Halpin, H. A., Morales-Suárez-Varela, M. M., & Martin-Moreno, J. M. (2010). Chronic disease prevention and the new public health. *Public Health Reviews*, 32(1), 120–154.
- [67]Hassan, N. M. & others. (2016). Work-life balance: A concern on effectiveness of job role and employee's engagement towards organizational goal. *International Journal of Academic Research in Business and Social Sciences*, 6(11), 411–422.
- [68]Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, 1017–1098.
- [69]Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies*, 64(4), 605–654. <https://doi.org/10.2307/2971733>
- [70]Heger, D. (2014). Work and well-being of informal caregivers in Europe. *Ruhr Economic Paper*, 512.
- [71]Heitmueller, A. (2007). The chicken or the egg?: Endogeneity in labour market participation of informal carers in England. *Journal of Health Economics*, 26(3), 536–559.
- [72]Henz, U. (2004). The effects of informal care on paid-work participation in Great Britain: A lifecourse perspective. *Ageing and Society*, 24, 851.
- [73]Hirazawa, M., & Yakita, A. (2017). Labor supply of elderly people, fertility, and economic development. *Journal of Macroeconomics*, 51, 75–96.
- [74]Huang, F., & Fu, W. (2017). Government purchase or family care? An empirical analysis based on the substitution effect of family care (in Chinese). *Nankai Economic Research*, 01, 136–152.
- [75]Human Resources and Social Security Ministry of China. (2016). *Guiding Opinions on the Pilot Program of Long-term Care Insurance System*.

- [76]Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 1–24.
- [77]Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4), 309.
- [78]Indicators, O., & Hagv íar, O. (2019). *Health at a glance 2019: OECD indicators*. OECD Publishing. Paris.
- [79]Insurance Association of China. (2019). *2018-2019 China Long-term Care Survey Report*.
- [80]Japan, S. (2016). Japan statistical yearbook 2016. Retrieved June, 9(2017), 1431–02.
- [81]Jaskiewicz, W., & Tulenko, K. (2012). Increasing community health worker productivity and effectiveness: A review of the influence of the work environment. *Human Resources for Health*, 10(1), 1–9.
- [82]Jeon, B., & Kwon, S. (2017). Health and long-term care systems for older people in the republic of Korea: Policy challenges and lessons. *Health Systems & Reform*, 3(3), 214–223.
- [83]Jeon, S.-H., & Pohl, R. V. (2017). Health and work in the family: Evidence from spouses' cancer diagnoses. *Journal of Health Economics*, 52, 1–18.
- [84]Jo, B. (2008). Causal inference in randomized experiments with mediational processes. *Psychological Methods*, 13(4), 314.
- [85]Johnson, R. W., & Sasso, A. T. L. (2000). *The trade-off between hours of paid employment and time assistance to elderly parents at midlife*.
- [86]Johnson, R. W., & Wiener, J. M. (2006). *A profile of frail older Americans and their caregivers*. Urban Institute Washington, DC.
- [87]Karlson, Kristian B, Holm, A., & Breen, R. (2010). Total, direct, and indirect effects in logit models. *Aarhus, Denmark: Centre for Strategic Research in Education, Aarhus University*.
- [88]Karlson, Kristian Bernt, & Holm, A. (2011). Decomposing primary and secondary effects: A new decomposition method. *Research in Social Stratification and Mobility*, 29(2), 221–237.
- [89]Kim, H. B., & Lim, W. (2015). Long-term care insurance, informal care, and medical expenditures. *Journal of Public Economics*, 125, 128–142.
- [90]Knutson, K. L., Phelan, J., Paskow, M. J., Roach, A., Whiton, K., Langer, G., Hillygus, D. S., Mokrzycki, M., Broughton, W. A., Chokroverty, S., & others.

- (2017). The national sleep foundation's sleep health index. *Sleep Health*, 3(4), 234–240.
- [91]Kohl, N. M., Mossakowski, K. N., Sanidad, I. I., Bird, O. T., & Nitz, L. H. (2019). Does the health of adult child caregivers vary by employment status in the United States? *Journal of Aging and Health*, 31(9), 1631–1651.
- [92]Kohler, U., Karlson, K. B., & Holm, A. (2011). Comparing coefficients of nested nonlinear probability models. *The Stata Journal*, 11(3), 420–438.
- [93]Kumamoto, K., Arai, Y., & Zarit, S. H. (2006). Use of home care services effectively reduces feelings of burden among family caregivers of disabled elderly in Japan: Preliminary results. *International Journal of Geriatric Psychiatry: A Journal of the Psychiatry of Late Life and Allied Sciences*, 21(2), 163–170.
- [94]Langley, P., Müller-Schwefe, G., Nicolaou, A., Liedgens, H., Pergolizzi, J., & Varrassi, G. (2010). The impact of pain on labor participation, absenteeism and presenteeism in the European Union. *Journal of Medical Economics*, 13(4), 662–672.
- [95]Lee, S.-H., Ogawa, N., & Matsukura, R. (2016). Japan's pension reform, labor market responses, and savings. *The Journal of the Economics of Ageing*, 8, 67–75.
- [96]Leigh, A. (2010). Informal care and labor market participation. *Labour Economics*, 17(1), 140–149.
- [97]Levine, C., Halper, D., Peist, A., & Gould, D. A. (2010). Bridging troubled waters: Family caregivers, transitions, and long-term care. *Health Affairs*, 29(1), 116–124.
- [98]Li, B.-R. (2018). A comparative study of long-term-care insurance for the elderly in qingdao, nantong and shanghai. *4th Annual International Conference on Management, Economics and Social Development (ICMESD 2018)*, 139–144.
- [99]Li, Q., & Tan, N. (2019). A review of research on the relationship between health and the labor supply of the elderly (in Chinese). *Journal of University of Electronic Science and Technology of China (Social Science Edition)*, 021(003), 38–47.
- [100] Lilly, M. B., Laporte, A., & Coyte, P. C. (2007). Labor market work and home care's unpaid caregivers: A systematic review of labor participation rates, predictors of labor market withdrawal, and hours of work. *The Milbank Quarterly*,

85(4), 641–690.

- [101] Lilly, M. B., Laporte, A., & Coyte, P. C. (2010). Do they care too much to work? The influence of caregiving intensity on the labour force participation of unpaid caregivers in Canada. *Journal of Health Economics*, 29(6), 895–903.
- [102] Lipszyc, B., Sail, E., Xavier, A., & others. (2012). *Long-term care: Need, use and expenditure in the EU-27*. Directorate General Economic and Financial Affairs (DG ECFIN), European
- [103] Liu, Y., Wei, Z., & Xie, F. (2014). Do women directors improve firm performance in China? *Journal of Corporate Finance*, 28, 169–184.
- [104] Lu, B., Mi, H., Zhu, Y., & Piggott, J. (2017). A sustainable long-term health care system for aging China: A case study of regional practice. *Health Systems & Reform*, 3(3), 182–190.
- [105] Lundberg, S. (1985). The added worker effect. *Journal of Labor Economics*, 3(1, Part 1), 11–37.
- [106] Ma, C., Yu, Q., Song, Z., & Chen, H. (2019). Long-term care insurance, the control of medical expenses and “value-based health care.” *China Industr. Econ*, 12, 42–59.
- [107] MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, Vol 62, 58, 593–614.
- [108] MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7(1), 83.
- [109] Makita, M. (2010). Gender roles and social policy in an ageing society. *International Journal of Ageing and Later Life*, 5(1), 77–106.
- [110] Mandel, H., & Semyonov, M. (2014). Gender pay gap and employment sector: Sources of earnings disparities in the United States, 1970–2010. *Demography*, 51(5), 1597–1618.
- [111] McGarry, K. (2007). 9. Does caregiving affect work? Evidence based on prior labor force experience. In *Health care issues in the United States and Japan* (pp. 209–228). University of Chicago Press.
- [112] McGarry, K. M. (2006). Does caregiving affect work? Evidence based on prior labor force experience. In *Health care issues in the United States and Japan* (pp. 209–228). University of Chicago Press.
- [113] Mellor, J. M. (2001). Long-term care and nursing home coverage: Are adult

- children substitutes for insurance policies? *Journal of Health Economics*, 20(4), 527–547.
- [114] Meng, D., Xu, G., He, L., Zhang, M., & Lin, D. (2017). What determines the preference for future living arrangements of middle-aged and older people in urban China? *Plos One*, 12(7), e0180764.
- [115] Meng, Q., & Tang, S. (2010). Universal coverage of health care in China: Challenges and opportunities. *World Health Report*.
- [116] Mentzakis, E., McNamee, P., & Ryan, M. (2009). Who cares and how much: Exploring the determinants of co-residential informal care. *Review of Economics of the Household*, 7(3), 283–303. <https://doi.org/10.1007/s11150-008-9047-0>
- [117] Michaud, P.-C., & Van Soest, A. (2008). Health and wealth of elderly couples: Causality tests using dynamic panel data models. *Journal of Health Economics*, 27(5), 1312–1325.
- [118] Michaud, P.-C., & Vermeulen, F. (2011). A collective labor supply model with complementarities in leisure: Identification and estimation by means of panel data. *Labour Economics*, 18(2), 159–167.
- [119] Moser, P., & Voena, A. (2012). Compulsory licensing: Evidence from the trading with the enemy act. *American Economic Review*, 102(1), 396–427.
- [120] Niimi, Y. (2018). Does providing informal elderly care hasten retirement? Evidence from Japan. *Review of Development Economics*, 22(3), 1039–1062.
- [121] Noon, M., Blyton, P., & Morrell, K. (2013). *The realities of work: Experiencing work and employment in contemporary society*. Macmillan International Higher Education.
- [122] Ojeda, V. D., Frank, R. G., McGuire, T. G., & Gilmer, T. P. (2010). Mental illness, nativity, gender and labor supply. *Health Economics*, 19(4), 396–421.
- [123] Oliveira Martins, J., & de la Maisonneuve, C. (2015). The future of health and long-term care spending. *OECD Journal: Economic Studies*, 2014(1), 61–96. https://doi.org/10.1787/eco_studies-2014-5jz0v44s66nw
- [124] Ozturk, Y., & Kose, T. (2019). Health, time allocation and work: Empirical evidence from Turkey. *Applied Economics*, 51(51), 5609–5622.
- [125] Pan, W., & Bai, H. (2015). *Propensity score analysis: Fundamentals and developments*. Guilford Publications.
- [126] Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3),

385–401.

- [127] Rellstab, S., Bakx, P., Garc ía-Gómez, P., & Van Doorslaer, E. (2020). The kids are alright-labour market effects of unexpected parental hospitalisations in the Netherlands. *Journal of Health Economics*, *69*, 102275.
- [128] Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, 143–155.
- [129] Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.
- [130] Rosenzweig, R. (1985). *Eight hours for what we will: Workers and leisure in an industrial city, 1870-1920*. Cambridge University Press.
- [131] Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3), 169–188.
- [132] Ryan, L. H., Smith, J., Antonucci, T. C., & Jackson, J. S. (2012). Cohort differences in the availability of informal caregivers: Are the Boomers at risk? *The Gerontologist*, *52*(2), 177–188.
- [133] Sageer, A., Rafat, S., & Agarwal, P. (2012). Identification of variables affecting employee satisfaction and their impact on the organization. *IOSR Journal of Business and Management*, *5*(1), 32–39.
- [134] Schmitz, H., & Westphal, M. (2015). Short-and medium-term effects of informal care provision on female caregivers' health. *Journal of Health Economics*, *42*, 174–185.
- [135] Schmitz, H., & Westphal, M. (2017). Informal care and long-term labor market outcomes. *Journal of Health Economics*, *56*, 1–18.
- [136] Schmitz, L. L. (2016). Do working conditions at older ages shape the health gradient? *Journal of Health Economics*, *50*, 183–197.
- [137] Schüz, B., Czerniawski, A., Davie, N., Miller, L., Quinn, M. G., King, C., Carr, A., Elliott, K.-E. J., Robinson, A., & Scott, J. L. (2015). Leisure time activities and mental health in informal dementia caregivers. *Applied Psychology: Health and Well-Being*, *7*(2), 230–248.
- [138] Skira, M. M. (2015). Dynamic wage and employment effects of elder parent care. *International Economic Review*, *56*(1), 63–93.
- [139] Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290–312.

- [140] Stabile, M., Laporte, A., & Coyte, P. C. (2006). Household responses to public home care programs. *Journal of Health Economics*, 25(4), 674–701.
- [141] State Council of China. (1994). *People's Republic of China Labor Law*.
- [142] Statistics, B. of L. (2016). Labor force statistics from the current population survey. *Bureau of Labor Statistics*.
- [143] Stephens, Jr, M. (2002). Worker displacement and the added worker effect. *Journal of Labor Economics*, 20(3), 504–537.
- [144] Strumpf, E. (2011). Medicaid's effect on single women's labor supply: Evidence from the introduction of Medicaid. *Journal of Health Economics*, 30(3), 531–548.
- [145] Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, 25(1), 1.
- [146] Sugawara, S., & Nakamura, J. (2014). Can formal elderly care stimulate female labor supply? The Japanese experience. *Journal of the Japanese and International Economies*, 34, 98–115.
- [147] Sugisawa, A., Sugisawa, H., Nakatani, Y., & Shibata, H. (1997). Effect of retirement on mental health and social well-being among elderly Japanese. [*Nihon Kosshu Eisei Zasshi*] *Japanese Journal of Public Health*, 44(2), 123–130.
- [148] Thompson, C. A., & Prottas, D. J. (2006). Relationships among organizational family support, job autonomy, perceived control, and employee well-being. *Journal of Occupational Health Psychology*, 11(1), 100.
- [149] Twenge, J. M. (2010). A review of the empirical evidence on generational differences in work attitudes. *Journal of Business and Psychology*, 25(2), 201–210.
- [150] Van Houtven, C. H., Coe, N. B., & Skira, M. M. (2013). The effect of informal care on work and wages. *Journal of Health Economics*, 32(1), 240–252.
- [151] Wang, Q., Zhou, Y., Ding, X., & Ying, X. (2018). Demand for long-term care insurance in China. *International Journal of Environmental Research and Public Health*, 15(1), 6.
- [152] Waxman, A. (2004). WHO global strategy on diet, physical activity and health. *Food and Nutrition Bulletin*, 25(3), 292–302.
- [153] White-Means, S. I., & Chang, C. F. (1994). Informal caregivers' leisure time and stress. *Journal of Family and Economic Issues*, 15(2), 117–136.

- [154] Wolf, D. A., & Soldo, B. J. (1994). Married women's allocation of time to employment and care of elderly parents. *Journal of Human Resources*, 1259–1276.
- [155] Yamada, H., & Shimizutani, S. (2015). Labor market outcomes of informal care provision in Japan. *The Journal of the Economics of Ageing*, 6, 79–88.
- [156] Yang, J., Wang, S., & DU, S. (2018). Regional comparison and implications of China's long-term care insurance system. *Chinese Journal of Health Policy*, 11(4), 1–7.
- [157] Yang, W., Jingwei He, A., Fang, L., & Mossialos, E. (2016). Financing institutional long-term care for the elderly in China: A policy evaluation of new models. *Health Policy and Planning*, 31(10), 1391–1401.
- [158] Yıkılkan, H., Aypak, C., & Görpelioglu, S. (2014). Depression, anxiety and quality of life in caregivers of long-term home care patients. *Archives of Psychiatric Nursing*, 28(3), 193–196.
- [159] Yu, Y., & Feng, J. (2018). The impact of family care on the utilization of medical services for the elderly (in Chinese). *Economics (Quarterly)*, 17(69(03)), 56–81.
- [160] Zhang, X., Zhao, X., & Harris, A. (2009). Chronic diseases and labour force participation in Australia. *Journal of Health Economics*, 28(1), 91–108.
- [161] Zhang, Y., & Yu, X. (2019). Evaluation of long-term care insurance policy in Chinese pilot cities. *International Journal of Environmental Research and Public Health*, 16(20), 3826.
- [162] Zhao, Y., Hu, Y., Smith, J. P., Strauss, J., & Yang, G. (2014). Cohort profile: The China health and retirement longitudinal study (CHARLS). *International Journal of Epidemiology*, 43(1), 61–68.
- [163] Zhou, Y., & Cao, R. (2017). The impact of the new rural insurance on the labor supply behavior of rural middle-aged and elderly people: A study based on the PSM-DID method (in Chinese). *Population and Economy*, 000(005), 97–107.
- [164] Zhu, Y., & Österle, A. (2019). China's policy experimentation on long-term care insurance: Implications for access. *The International Journal of Health Planning and Management*, 34(4), e1661–e1674.

Appendix A. Key Features of China's LTCI Pilots as of April 2021

[Insert Table A.1 here]

[Insert Figure A.1 here]

[Insert Table A.2 here]

[Insert Table A.3 here]