Maternity Nonagricultural Employment, Rural-Urban Migration, Maternal Care and Child Health – Evidence from Rural China

Jianfeng Yao

(Preliminary and Incomplete)
Abstract:

The ratio of women participating in nonagricultural work has increased significantly in rural China since the so-called “Market Economy” was adopted in the 1980s. This phenomenon has had a great impact on family income earning patterns and poverty reduction. However, from the perspective of the woman’s dual identity in the family – income earner and child caregiver – the change in work patterns and labor supply behavior has had an uncertain effect on child welfare and human capital accumulation. This paper studies maternal labor supply patterns for those female “peasant workers” in rural areas and discusses women’s labor participation behavior and its effect on child health outcomes. I seek to develop and estimate a dynamic model of the agricultural, nonagricultural and migratory employment and maternal care decisions of rural women to evaluate the effects of these choices on child health. Other effects such as maternal agricultural and nonagricultural working time and maternal income effects will also be examined. Children with different genders may be affected differently; girls in general have a lower health state than boys. Policy implication strategies will also be discussed.
Introduction:

In many developing countries, women in the family usually have two folded identities – income earner, and child caregiver. Dismissing either one in the policy making process will have an unexpected impact on the implementation and effectiveness of a policy. Recently, more attention has been paid to rural woman’s dual identity. This is partly because woman’s labor force participation rate in developing countries has increased dramatically since the 1980s (Mehra and Gammage, 1999), and partly because the change in economic structure brought by development within a country may alter the characteristics of woman’s work. For example, agricultural work is gradually replaced by jobs in manufacture, service and commercial sectors (Glick, 2002). These changes may have influential effects on rural women and their family lifestyles through changes in maternal income, employment flexibility, maternal care demand and choices. This issue has earned a special meaning for China after its economic booms.

In recent years, China has gone through a dramatic economic growth and structural change. The most noticeable impact of this change has been the transformation of the agricultural labor force (Zhang, et. al, 2004). The emergence and development of rural nonagricultural labor markets has created many job opportunities (Meng, 2000). At same time, the nonagricultural labor force participation rate has risen significantly (Lohmar, et al., 2001; Rozelle, et al., 1999). These factors significantly changed job characteristics and opportunities facing women in rural areas. Many recent studies have looked at rural labor market development and nonagricultural labor migration. However, most of these studies have only focused on the employment behavior of the labor force. As for the dual identity of women in rural areas, more attention is paid to its effect on employment, dismissing its impact on human capital accumulation brought by child caregivers. This may lead to a bias in policy evaluation. When analyzing woman’s labor participation behaviors, ignoring the important role played by the childcare giver will likely compromise the explanatory power of the research results.

As the child care giver, the mother’s employment transition behavior will strongly affect the welfare of the next generation. This impact is closely related to the two-fold identity of woman within the family. From the perspective of economic income, an increase in female job opportunities, especially nonagricultural job opportunities, is favorable for increasing household income and reducing poverty. Mothers usually have strong
preferences on investment in the childcare-related commodities (Glick, 2002), which will undoubtedly increase child welfare. Further, an increase in the opportunity cost associated with raising a child will likely reduce fertility rates, and therefore lead to a higher investment in the quality of child care (Becker and Lewis, 1973). An increase in job opportunities outside of the family will also likely to improve woman’s economic status, and therefore stimulate the investment in female children (Hare, 1999).

Alternatively, as primary childcare givers within the family, mothers face time constraints that generate potential conflicts between market employment behavior and behavior in the family. Working mothers may lack enough maternal care time, and care quality might be lowered because of working stress and exhaustion (Desai, et al., 1989). Even with market or nonmarket child care substitutes, such as babysitters, relatives or older children, problems like high cost and low quality may well exist (Glick and Sahn, 1998; McGuire and Popkin, 1990). For fathers in the family, the impact of their employment on child welfare is mainly through their income. In many developing countries, fathers in the family engage less in childcare activities than do mothers, especially for young children (Evans, 1995). From this point of view, the impact of woman’s labor supply on child development is far more complicated.

Therefore, although nonagricultural employment plays an important role in increasing “peasant worker” income, when we take into account its negative effect on the quantity and quality of childcare brought by the increase in working time, migratory nonagricultural employment also gives rise to a special group of children called “left behind children.” As a result, nonagricultural employment, especially rural-urban migratory employment, still has ambiguous effects on the next generation’s welfare and human capital accumulation in rural areas. Therefore, theoretical and empirical studies should be conducted carefully with respect to this issue. Studies of the effect of rural mother’s labor supply behavior on their child’s human capital accumulation (including health, education, etc.) have important policy implications. This paper mainly focuses on the impact of rural woman’s labor participation choices and maternal care decisions on child health. The data are from the China Health and Nutrition Survey (CHNS), and I define a sickness frequency dummy as the child health index. By categorizing different types and places of work, I seek to investigate different working decisions and maternal care choice’s effect on child health. Many recent literatures only concentrate on whether a rural mother works or not, and dismisses the difference between self-employers (including agricultural work) and those who do not participate in any working activities. As a consequence, the explanatory power of their studies
Background:

Maternal Employment on Child Health:

The extensive literature studying the effect of maternal employment on children has primarily focused on child cognitive development as an outcome, perhaps because of the wide availability of objective measures such as academic performance, and on children at early ages (Bernal, 2008; Blau and Grossberg, 1992; Desai, et al., 1989; Kaestner and Corman, 1995; Ruhm, 2004; Waldfogel, et al., 2002). The findings are mixed, but generally the estimated effect of maternal employment is small. The role of maternal employment in a child’s non-cognitive development or physical conditions (such as health outcomes) in developing countries has received comparatively scant attention. There is a growing recent literature that argues maternal employment increases the risk of childhood obesity, though only for the higher socioeconomic status populations (Anderson, et al., 2003). Gordon, et al. (2007) use a fixed effects strategy to measure the effects of maternal employment (and child care) on child injuries and infectious diseases for children ages 12 to 36 months.

Several recent studies have sought to analyze the effects of maternal employment on the health of children. Gennetian, et al. (2010) identify the effects of low-income mothers' employment on the health of young children by exploiting a welfare-to-work experiment, the National Evaluation of Welfare-to-Work Strategies (NEWWS). They find that among the low-income children in the sample, maternal employment decreases a child's probability of being in good or excellent health by a modest amount. Ruhm (2008) uses the National Longitudinal Survey of Youth (NLSY) to analyze the effect of maternal employment on a cohort of children ages 10 and 11. He employs a fixed effects strategy to control for fixed mother and family characteristics. He finds large differences in effects by the child's socioeconomic status (and other proxies for disadvantage), where disadvantaged children see no effect or benefit from maternal employment and advantaged children experience harmful consequences.
Baker, et al. (2008) estimate the effect of maternal labor supply on young children's health by examining the impact of a local childcare subsidy program in Quebec in the late 1990s. They use a difference-in-differences identification strategy and conclude that the policy led to an increase in maternal labor supply, an increase in formal childcare enrollment, and a decline in health for children. The authors consider the impact of the childcare subsidy program on the child who is eligible and therefore cannot separate the direct effect of childcare from the effect of maternal employment.

Liu (2008) analyzes the effects of maternal labor supply on children’s health in rural China. With the data from the China Health and Nutrition Survey (CHNS) and a static model, he finds that an increase in mother’s working hours reduces her time available for children and consequently has a negative effect on child health, while the additional income generated contributes to improving child’s nutritional status. With the labor working time effect outweighing the income effect, maternal work overall shows a small negative effect on child health, and the effect is more pronounced for nonagricultural work than agricultural activities. These results suggest that economic structural changes have intensified the conflict of women’s dual identity as both income earners and as caregivers.

**Work related migration on left-behind children:**

With a growing number of children left-behind living in under-developed rural areas and being separated from one or two of their parents, there will likely be impacts on these children’s physical, psychological, behavioral and health conditions. Left-behind children are usually taken care of by their relatives or grandparents. There are two different viewpoints with respect to this issue. The first group of researchers argues that, under the condition of unavailability of parental care, grandparents may provide basic family love and security that is more favorable to child development (Bert, et al., 2005). Solomon, et al. (2005) conclude that there is no significant difference between the health status of children raised by their relatives and those raised by their biological parents from the representative data from the United States. Other researchers, however, argue that most grandparents are relatively advanced in age, and have low education and poor health. Therefore, they are not always able to provide enough resources for childcare, and the child’s physical condition is likely to be negatively affected (Gaudin, et al., 1993; Kirby, et al., 2002). Matthew, et al. find that children who live with their parents are generally healthier after controlling for individual socioeconomic factors.
There have been relatively fewer econometric analyses in the studies conducted by Chinese researchers. Although most studies argue for an ambiguous effect of family structure on child health, farmers in underdeveloped villages have lower income and lack access to quality medical facilities, therefore, the health condition of children left behind is more vulnerable to family structure. Using CHNS data from 2000 to 2006, Chen (2009) analyzes the health status of the left-behind children in rural China, based on the height-for-age Z score (HAZ) and body mass index (BMI). The results show that the left-behind children from zero to five years old have the same health status as those living with two biological parents whose health is dependent on household income and access to medical systems. However, the left-behind children from six to eighteen years old have poorer health than those living with their parents. Focusing on the left-behind children over five years old and living without mothers, he further finds that the negative effect of a shortage of mother’s care on child health is pronounced not only for low income families but also for high income families. Therefore, it is important for adolescent development to increase household income and maternal care time to improve the nutrition of children.

Through measuring the effect of child care quality, Currie and Hotz (2004) suggest an important role for supervision in avoiding accidents and injuries among young children. They find that the incidence of unintentional injury for children under age 5 is reduced in states with more stringent childcare regulations. In a related work, Aizer (2004) shows that after-school supervision of adolescents ages 10 to 14 has a strong effect on their well-being as measured by criminal activity and behavior problems. Aizer uses a sample from the National Longitudinal Survey of Youth (NLSY) to estimate several fixed effects models using variation in supervision between and within families. If children whose mothers’ work spend more time unsupervised, those children may have a higher risk of accident or injury, which may also lead to additional hospitalizations.

**Theoretical framework:**

In this section, I present a theoretical model of married rural mothers’ decisions regarding agricultural, nonagricultural, and migratory nonagricultural work, maternal care time and child health expenditures, and identify how these decisions may affect child health. A woman makes choices on different types and places of work and maternal care in each period $t$ following child age from 3 to 17. I also assume that each family has
only one child because of the one child policy imposed by the Chinese government.

The timing of the model is as follows: at the beginning of $t$, a mother makes an employment and corresponding migration decision based on the expected utility of different types of work and associated workplaces. Each type of work is characterized by a specific location of work. As long as a mother chooses agricultural work, she can only engage in such work in the village. However, if she only chooses nonagricultural work, she then has the option of working locally or leaving for cities and leaving her child behind. The choice of employment and migration in each period will depend on the mother’s wages, working time and home production distribution. After choosing a type of work and workplace, the mother’s wage offer and working time can be observed. Further assuming that the mother always accepts earning offers in each period for low-skill and low-income jobs, she then chooses how much maternal care time and health care expenditures are spent on the child. Given the mother’s choices and inputs, a child’s health status will finally be achieved.

I allow for first stage employment and migration choices ($a = 0, 1, 2, 3(0), 3(1)$) as no work, agricultural work in the village, rural agricultural work, and nonagricultural work in the village. Conditional on nonagricultural work, a woman may also choose to work in the village or city. The second stage choice ($m = 0, 1, 2$) is denoted as spending no time, some time or enough time on maternal care (I discretize maternal care time into 3 categories). The choice of health care expenditures on a child ($e = 0, 1, 2$) is also discretized into 3 categories (low, medium or high). This means in the second stage there are $3 \times 3$ possible options in a woman’s choice set $\Gamma_2$ with 9 alternatives. I denote $d_{a,t}^1$ and $d_{j,t}^2$ as indicator functions that equal 1 if an alternative is chosen in the first and second stage at time $t$.

**Utility Function:**

The current utility function given the mother’s choice is given by:

$$U_t = U_t(x_t, S_t, h_t, m_t, a_t; CM_t, R_t) = \frac{x_t^{\rho_1}}{\rho_1} \left[ 1 + \alpha_1 (h_t > 0) + \alpha_2 (m_t > 0) \right] + \frac{\alpha_3 (S_t + \rho_2)^{\rho_3}}{\rho_3}$$

$$+ \alpha_4 h_t + \alpha_5 m_t + \alpha_6 h_m_t + CM_t (\omega_1 + \omega_2 R_t) + \sum_{j=0}^{J} d_{j,t}^2 \varepsilon_{jt} + \varepsilon_{at}$$

(1)

And consumption $x_t$ is given by the budget constraint:
Each individual is endowed with a fixed amount of time that she uses to provide maternal care, work and leisure:
\[ T = m_t + h_t + l_t \] (3)

Here, is the observed binary variable defined below, indicating whether the child has been sick within past three weeks. is mother’s hourly wage and is father’s average yearly income, which is assumed to be exogenously determined. is the health expenditure on the child, such as spending on appropriate nutrition and vaccinations that affect the child’s morbidity. is the home production function compensation for no work.

The utility function is has the common CRRA form in consumption. represents the indicator function equal to 1 if the argument is true and 0 otherwise. The parameter and are the disutilities from working and time spent on maternal care, and indicates that the extra disutility from providing maternal care is derived from working. The mother also obtains a disutility from child’s sickness condition according to the CRRA function with parameter . Community level attributes and sociodemographic features of the mother also enter into the utility function. are maternal evaluations of unobserved attributes of five employment and migration states, and are unobserved attributes of maternal care time and child health expenditures. These random preference shocks are assumed to be independently and identically distributed over different alternatives across individuals and time periods with an extreme value distribution, with variance parameters and respectively.

**Probability of sickness:**

Health status is not observed and is uncertain in the next decision horizon. Therefore, we may use weight-for-age z-score (WAZ), body mass index (BMI) or a sickness frequency indicator to reveal a child’s
health condition. The latent health status at age $t+1$, $health_{t+1}$, depends on child age and gender, present child health expenditure, mother’s maternal care time and working time (affecting the quality of child care), mother characteristics, some community level exogenous variables and mother health status at $t$. Define:

$$health_{t+1} = \beta_1 age_{t+1} + \beta_2 e_t + \beta_3 edu_{t+1} + \beta_4 m_{t+1} + \beta_5 CM_{t+1} + \beta_6 Z_{t+1} + \beta_7 h_{t+1} + \beta_8 (1 - d_{3(t+1)}) + \beta_9 gender + S_t (\beta_{10} + \beta_{11} sh_t) + \mu^m + \epsilon_{health_{t+1}} \quad \text{(4)}$$

Where $\epsilon_{health_{t+1}}$ is the health shock that follows the serially independent standard normal distribution, and $\mu^m$ is the unobserved trait of the child. $sh_t$ is defined as the duration of sickness (number of sick years up to $t$). If the child is not frequently sick at age $t$ ($S_t = 0$), all previous sick years will have no impact on health status at age $t+1$. However, if the child is frequently sick ($S_t = 1$), the number of continuous frequent sick years up to $t, sh_t$, will affect health status at age $t+1$. This will be demonstrated in detail in the discussion of the evolution of state variables. Then we may have:

$$S_{t+1} = 1 \text{ if } health_{t+1} > 0 \text{ (sick)} \text{ and } S_{t+1} = 0 \text{ if } health_{t+1} \leq 0 \text{ (not sick)}.$$ 

Note that the $health_{t+1}$ here is related to, but different from, the health capital of Grossman (1972). To model Grossman’s health capital, $health_{t+1}$, would need to depend on $health_t$. Since both $health_{t+1}$ and $health_t$ are not observed, such a model would be problematic to estimate. Instead, we may use an observed binary variable $S_t$ and a cumulative stock variable $sh_t$ to approximate $health_t$. Compared with the effect of the most current health status $S_t$, the effect of $sh_t$ might be small. The parameter $\beta_8$ reflects the idea that more educated children may have a better knowledge of health issues and thereby refrain from activities that are harmful to their health. $\beta_8$ represents the psychological and physical effect on the child when the mother is not in town. Mother’s characteristics $Z_{t+1}$ such as education and age may also play an important role in determining child health.

**Mother’s wage:**

I do not model father’s wage; instead I assume father’s wage is exogenously determined. I further assume that mother’s wage is a logarithm function of her characteristics such as educational attainment, $edu_t$, and
age, age, her skill type k, and her work experience, ep\. Since a woman is not able to choose part-time work, her child’s health is assumed to have an impact on her wage or work efficiency. \( \tau_{pt} \) is the labor market condition and job attributes at t in the city or village. Therefore, the wage function can be defined as:

\[
\ln w_t = r_{1} edu + r_{2} ep_{t} + r_{3} ep_{t}^{2} + r_{4} age_{t} + S_{i}(r_{5} + r_{6} sh_{t}) \\
+ r_{7} 1(ep_{t} = ep_{t-1}) + r_{8} \tau_{pt} + \sum_{a=1}^{3(1)} r_{9a} d_{a}^{1} + \sum_{k=1}^{2} r_{ok} 1(skilltype = k) + \epsilon_{wt}
\]

where \( r_{7} \) represents the adjustment cost if the individual did not work in the previous period. \( \epsilon_{wt} \) is the idiosyncratic wage shock. Mother’s skill endowments (k =1 when the skill endowment is low, k=2 when the skill endowment is high) is not observed. Skill types enter into the wage function to reflect the effect of different market skills on wages. I will refer to this in detail in the data simulation section later. Moreover, being employed in the agricultural sector versus nonagricultural sectors indicates different wages.

**Mother’s working time:**

Working time is assumed to be not chosen by the individual as part-time or full-time for low-income jobs that most farmer workers choose, and therefore becomes endogenous just like wage:

\[
h_t = \phi_{1} \tau_{pt} + \sum_{a=1}^{3(1)} d_{a}^{1} \left[ \phi_{2a} + \sum_{q=1}^{O} \phi_{aq} 1(worktype = q) \right] + \epsilon_{ht}
\]

Mother’s working time is mainly dependent on the type of work she chooses (agriculture or nonagricultural) and the specific type of work within each sector that can be observed from the data. Therefore, unlike mother’s skill type endowment, work type can be observed. Again, the labor market condition and job attributes at t in the city or village, \( \tau_{pt} \), is assumed to affect working time. \( \epsilon_{ht} \) is the idiosyncratic working time shock.

**Home production function:**

The output of home production cannot be unobservable to econometricians, but is observed to the individuals. Any output that the individual produces to lower household expenditures, and any compensation that she may receive when staying at home, are included as the output of home production. For simplicity, the home production function is assumed to only depend on child health at home:

\[
pr_{t} = \bar{e} + S_{i}(\phi_{1} + \phi_{2} sh_{t}) + \epsilon_{pt}
\]
where $\tilde{e}$ is constant and $\varphi_1$ and $\varphi_2$ are coefficients of health status. The shock to wage equation, $\varepsilon_{wt}$, working time equation, $\varepsilon_{ht}$, and home production, $\varepsilon_{pt}$, are serially independent and follow normal distribution with specific variances and covariance.

**Evolution of state variables:**

The state space of this dynamic programming model is:

$$\Omega_t = \{ep_t, S_t, sh_t, Z_t, age_t, edu_t, d_{t-1}^1, d_{t-1}^2, \varepsilon_{wt}, \varepsilon_{ht}, \varepsilon_{pt}\}$$

It is important to describe how the elements of the state space evolve. I will only focus on the first three variables in the state space. The evolution of the rest of the elements is either obvious or independent across years. Each woman and child has a set of individual specific variables that stay fixed over time, or that are assumed to evolve exogenously. These include children’s age and education, mother’s skill endowment, race, age, and education, and her spouse’s average income.

The law of motion of mother’s work experience is: $ep_{t+1} = ep_t + (1 - d_{h,t}^1)$. This is to say that mother’s work experience increases by one time period if and only if she works at child age $t$. The sickness frequency dummy at age $t+1$, $S_{t+1}$, takes value at the end of $t+1$, after the choices regarding child health expenditure and maternal care time at $t$ have been made ($S_{t+1} = 1$ if child is sick; $S_{t+1} = 0$ if the child is not sick). The variable $sh_t$ measures the duration of frequent sickness up to $t$ (not including age $t$). In particular, $sh_t$ takes the following form: $sh_{t+1} = S_t (sh_t + S_t)$, meaning that $sh_{t+1} = 0$ if $S_t = 0$. For instance, if the child is sick at age $t$ and age $t-1$, but not sick at age $t-2$, then $sh_{t+1} = 1$.

**Solution Method:**

The maximization problem is set into a dynamic programming framework. The value function can be written as the maximum over alternative specific value functions, each of which follows the Bellman equation.

**Expected utility with respect to each maternal care time and child health expenditure choice $j$:**
If the mother ends up with employment and workplace alternative $a$ at time $t$, with her realized wage, $w_t$, working time, $h_t$, husband's income, $y_t$, and labor market conditions and job attributes at her workplace, $\tau_{pt}$.

Learning the realization of each state and preference shocks as with each maternal care and health expenditure decision, the expected utility associated with each maternal care time and child health expenditure choice can be denoted as:

$$V_{at}^j (\Omega_t, \tau_{pt}, \mu, \varepsilon_j, \xi_{at}) = E[U_j (x_t, S_t, h_t, m_t, a_t) | \Omega_t, \tau_{pt}, \mu, \varepsilon_j, \xi_{at}]$$

$$= \frac{\alpha_2}{\rho_3} \sum_{s=0}^{1} \left\{ \Pr(S_t = s | \Omega_t, \mu, h_t, m_t, a_t)(S_t + \rho_2)^{\rho_1} \right\} + \overline{U}_j (x_t, h_t, m_t, a_t) + \varepsilon_j + \xi_{at}$$

where $\mu$ represents the vector of unobserved mother and child time invariant traits. The child’s health outcome is stochastic and realized after the maternal time and health expenditure decision is made. Therefore the calculation of expectation is required after observing mother’s wage, working time and preference shocks. The utility function is assumed to be separable additive with independence of $\varepsilon_j, \xi_{at}$ and $\varepsilon_{health}$, and $\overline{U}_j (x_t, h_t, m_t, a_t)$ is the deterministic utility function without the sickness dummy. This implies that, given the choice of employment and workplace alternative, at wage rate, $w_t$, and working time, $h_t$, labor market conditions and job attributes at her workplace, $\tau_{pt}$, the optimal decision of maternal care time and child health expenditure can be defined as:

$$f_t = \arg \max \left\{ \overline{V}_{at}^j (\Omega_t, \tau_{pt}, \mu) + \varepsilon_j \right\}$$

Assuming $\varepsilon_j$ is the independently and identically distributed extreme value error, the probability of choosing a specific maternal care time conditional on a set of state variables and location market conditions should be:

$$\Pr(f_t = j | \Omega_t, \tau_{pt}, \mu) = \frac{\exp[\overline{V}_{at}^j (\Omega_t, \tau_{pt}, \mu) / b_j]}{\sum_{j=0}^{J} \exp[\overline{V}_{at}^j (\Omega_t, \tau_{pt}, \mu) / b_j]}$$

**Expected maximum utility associated with choosing option $a$:**

In the first stage of decision-making for employment and workplace, the mother knows neither actual wage that will be offered nor actual working time that will be available to her. In addition, she does not know the
realization of her preference shocks associated with her maternal care time and child health expenditure decision, or child health outcome shocks. The expectation is respect to the extreme value distributed preference shocks, \( \varepsilon_j \), and \( b_j \) is the variance of these shocks. The optimization problem is subject to budget constraint and the terminal value function can be estimated by a polynomial function. Therefore, the expected maximum utility associated with choosing employment and workplace alternative can be expressed by:

\[
V_{at}(\Omega_i, \tau_{pt}, \mu, \xi_{at}) = E \max_j V_{at}^j(\Omega_i, \tau_{pt}, \mu, \varepsilon_j, \xi_{at}) = b_j \log \left( \sum_{j=0}^{J} \exp \left( V_{at}^j(\Omega_i, \tau_{pt}, \mu) / b_j \right) \right)
\]

**11**

\[+ \xi_{at} = \bar{V}_{at}(\Omega_i, \tau_{pt}, \mu) + \xi_{at}\]

Likewise, at the time of employment and workplace decision, \( \xi_{at} \) is known and the optimal decision for the mother can be defined as:

\[a_t = \arg \max \left\{ V_{at}(\Omega_i, \tau_{pt}, \mu) + \xi_{at} \right\}\]

After assuming the independently and identically distributed extreme value error for \( \xi_{at} \), the probability that a specific employment and workplace alternative is chosen can be:

\[
\Pr(a_t = a|\Omega_i, \tau_{pt}, \mu) = \frac{\exp[V_{at}(\Omega_i, \tau_{pt}, \mu) / b_a]}{\sum_{a=0}^{3} \exp[V_{at}(\Omega_i, \tau_{pt}, \mu) / b_a]}
\]

**12**

However, when \( a \) is chosen as 3, meaning only nonagricultural work is chosen by the mother, she will then be able to choose to either leave for the city or stay in the village to engage in such nonagricultural work. Therefore, the form of this probability can be used in the estimation of a nested logit model:

\[
\Pr(a_t = 3|\Omega_i, \tau_{pt}, \mu) = \frac{\exp[V_{at}(\Omega_i, \tau_{pt}, \mu) / b_a]}{\sum_{a=0}^{3} \exp[V_{at}(\Omega_i, \tau_{pt}, \mu) / b_a]}
\]

**13**

\[
\Pr(p_t = p|a_t = 3, \Omega_i, \tau_{pt}, \mu) = \frac{\exp[V_{at}(\Omega_i, \tau_{pt}, \mu) + b_a I_p]}{\sum_{p'=0}^{1} \exp[V_{at}(\Omega_i, \tau_{pt}, \mu) + b_a I_p]}
\]

**14**

Finally, the probability of nonagricultural work with a different workplace decision can be defined as:

\[
\Pr(a_t = 3(p)|\Omega_i, \tau_{pt}, \mu) = \Pr(a_t = 3|\Omega_i, \tau_{pt}, \mu) \cdot \Pr(p_t = p|a_t = 3, \Omega_i, \tau_{pt}, \mu)
\]

**15**
\[
I_a = \ln \left( \sum_{p=0}^{1} \exp \left[ \mathcal{P}_{p,t} (\Omega_t, \tau_{p,t}, \mu) \right] \right) \quad \text{and} \quad p = 0, 1
\]

**Likelihood contribution for mother \( i \):**

The CHNS data allow access to the longitudinal information on the mother’s agricultural, nonagricultural work and workplace decisions, mother’s wage and working time, maternal care time, child health expenditure and child sickness status. Home production when mother is not engaged in market work is not observed. Given optimal decision rules and specifying a continuous distribution for the unobserved heterogeneity factor \( \mu \) with density, the unconditional likelihood contribution for mother \( i \) is given by:

\[
L_i = \prod_{t=1}^{T} \left[ \prod_{a=0}^{2} \Pr (a_i = a \mid \Omega_t, \tau_{pt}, \mu) \prod_{p=0}^{1} \Pr (p_i = p \mid a_i = 3, \Omega_t, \tau_{pt}, \mu) \right] \times
\left\{ \left( \phi \left( W_i \mid \Omega_t, a_i, \tau_{pt}, \mu \right) \cdot \Phi \left( h_i \mid \Omega_t, \tau_{pt}, a_i, \mu \right) \right) \right\}^{1(\omega=0)} \cdot \left\{ \prod_{j} \Pr (j_i = j \mid \Omega_t, \tau_{pt}, a_i, \mu) \right\} \cdot \Pr (S_i = S \mid \Omega_t, \tau_{pt}, j, a_i, \mu) \bigg\} \mathcal{g} (\mu) d\mu
\]

The likelihood function for the sample is:

\[
L = \prod_{i=1}^{I} L_i
\]

**An alternative estimation method: Indirect Inference (II): Keane and Smith (2003):**

The approach of indirect inference first selects a simple descriptive statistical model, and indirectly matches the coefficient estimates of the descriptive model from the simulated data with those from the observed data. This method circumvents the need to construct the choice probabilities generated by the economic model, which sometimes causes the problem of high dimensional integration, especially when the number of possible outcome is large, because the model is not based on forming the likelihood for the model. The idea of indirect inference is to simulate data from the structural model of interest. A descriptive statistical model is then chosen to provide a rich description on the pattern of co-variation in the data. Loosely speaking, indirect inference generates an estimate \( \hat{\beta} \) of the structural parameters by choosing a set of structural parameters \( \beta \) so as to
make \( \hat{\theta} \) and \( \tilde{\theta}(\beta) \), the set of parameters that maximizes the likelihood function of the descriptive statistical model from the observed and simulated data, as close as possible. When generating simulated data sets, the set of random draws, \( \{\eta_u\} \), is held fixed for different values of \( \beta \). Also, from the Monte Carlo experiment, Keane and Smith’s (2003) general indirect inference approach (GII) takes approximately 30 percent less time than computing a set of estimates using a simulated maximum likelihood estimation approach. The application of GII can be implemented in four steps:

**Step 1: Descriptive statistical model estimation using the observed data:**

A linear probability model is chosen as an appropriate descriptive statistical model to provide a good description of the data, due to its computational tractability and statistical efficiency. Denote \( \{y_{it}, x_{it}\}_{i=1}^{N}, t = 1, \ldots, T \) as observed choices and outcomes for mother \( i \) at time \( t \). The observed choices include agricultural work and nonagricultural work at different places. The observed outcome contains mother’s wages, working time, home production and child healthy or sick status. Then, the descriptive statistical model is given by

\[
y_{it} = \gamma_i x_{it} + \eta_i, \; \eta_i \sim N(0, K_i)
\]

(17)

where \( x_{it} \) is a set of regressors that will be discussed in the next section, and \( \theta_i = (\eta_i, K_i) \) is the set of parameters to be estimated. Therefore, the likelihood function of this descriptive statistical model can be written as:

\[
L(y; z, \Theta) = \prod_{i=1}^{N} \prod_{t=1}^{T} l(y_{it}, x_{it}, \theta_t)
\]

(18)

where \( z \) is the initial exogenous variables that include child health or sick status, mother and children’s characteristics, and working experience, and \( \Theta \) represents the parameter set \( \{\theta_t\}_{t=1}^{T} \). Therefore, a set of parameters can be found to maximize the likelihood function of this descriptive statistical model as:

\[
\hat{\Theta} = \arg \max L(y; z, \Theta)
\]

**Step 2: Data Simulation from the structural model:**

Here I will simulate the mother’s choice from child age 3 to 17. From Keane and Smith (2003), and given a set of initial exogenous variables \( z \) and a set of parameters \( \beta \), it is possible to generate statistically independent
simulated data sets: \( \{ \tilde{y}_{it}^{m}(\beta) \}_{i=1}^{N}, m = 1, \ldots, M; t = 1, \ldots, T \). Here, \( M \) is the total number of data sets generated, and \( y_{it} \) and \( \tilde{y}_{it} \) contain the same type of choices and outcomes described in step 1. Each of these \( M \)-simulated datasets is generated using the same set of observed exogenous variables \( z \); however, the set of random error draws is held fixed for different values of \( \beta \).

**Step 3: Using the simulated data to estimate descriptive statistical model:**

The descriptive statistical model (17) can be estimated using the simulated data sets. However, according to Keane and Smith (2003), the simulated discrete variables cannot be directly used in the descriptive statistical model because of the non-smooth objective function. Therefore, they propose the idea of general indirect inference (GII) in which a series of functions of latent utility are used to substitute the simulated discrete choice variables \( \tilde{d}_{ia}^{1}(\beta) \) and \( \tilde{d}_{ij}^{2}(\beta) \). More specifically, \( \tilde{d}_{ia}^{1}(\beta) \) should be replaced by

\[
\tilde{d}_{ia}^{1}(\beta, \lambda) = \frac{\exp[\tilde{V}_{ai}(\Omega_{t}, \tau_{pt}, \mu; \beta)/\lambda]}{\sum_{a' = 0}^{1} \exp[\tilde{V}_{a'i}(\Omega_{t}, \tau_{pt}, \mu; \beta)/\lambda]} 
\]

and

\[
\frac{1}{\lambda} \sum_{a' = 0}^{1} \exp[\tilde{V}_{a'i}(\Omega_{t}, \tau_{pt}, \mu; \beta)/\lambda] \times \frac{1}{\lambda} \sum_{p' = 0}^{1} \exp[\tilde{V}_{a'i}(\Omega_{t}, \tau_{pt}, \mu; \beta) + \lambda I_{a}] 
\]

when \( a = 3 \)

\( \tilde{d}_{ij}^{2}(\beta) \) has to be replaced by

\[
\tilde{d}_{ij}^{2}(\beta, \lambda) = \frac{\exp[\tilde{V}_{ai}(\Omega_{t}, \tau_{pt}, \mu; \beta)/\lambda]}{\sum_{j' = 0}^{1} \exp[\tilde{V}_{ai}(\Omega_{t}, \tau_{pt}, \mu; \beta)/\lambda]} 
\]

where \( \lambda \) is the smooth parameter. Because the latent utilities are smooth functions of the parameter \( \beta \), \( \tilde{d}_{ia}^{1}(\beta, \lambda) \) and \( \tilde{d}_{ij}^{2}(\beta, \lambda) \) are also smooth functions of \( \beta \). In addition, as the smooth parameter approaches zero, both \( \tilde{d}_{ia}^{1}(\beta, \lambda) \) and \( \tilde{d}_{ij}^{2}(\beta, \lambda) \) approach 1 if an alternative has the highest utility, and approaches zero otherwise.

Mother’s wages and working time are observed only if she worked during that period. In order to match the simulated wage with the observed wage, we may set wage equal to zero for those who did not work during that
time period. Therefore, \( \tilde{w}_t(\beta) \) and \( \tilde{h}_t(\beta) \) should be replaced by

\[
\tilde{w}_t(\beta;\lambda) = \tilde{w}_t(\beta) \ast \left( \sum_{a=1}^{3(1)} \tilde{d}_{a,t}^1(\beta,\lambda) \right)
\]

and

\[
\tilde{h}_t(\beta;\lambda) = \tilde{h}_t(\beta) \ast \left( \sum_{a=1}^{3(1)} \tilde{d}_{a,t}^1(\beta,\lambda) \right)
\]

respectively. The estimated parameters of the descriptive statistical model that uses simulated data are smooth functions of \( \beta \), since \( \tilde{w}_t(\beta) \), \( \tilde{h}_t(\beta) \) and \( \left( \sum_{a=1}^{3(1)} \tilde{d}_{a,t}^1(\beta,\lambda) \right) \) are smooth functions of \( \beta \). When the smoothing parameter \( \lambda \) approaches zero, \( \tilde{w}_t(\beta) \), \( \tilde{h}_t(\beta) \) and \( \left( \sum_{a=1}^{3(1)} \tilde{d}_{a,t}^1(\beta,\lambda) \right) \) approach \( \tilde{w}_t(\beta) \) and \( \tilde{h}_t(\beta) \) respectively, if the chosen alternative has the highest latent utility, and approaches zero otherwise. The sickness dummy \( \tilde{S}_{r,t}(\beta) \) should be replaced by a continuous function

\[
\tilde{S}_{r,t}(\beta;\lambda) = \frac{\exp[Health_{r,t}(\beta)/\lambda]}{1 + \exp[Health_{r,t}(\beta)/\lambda]},
\]

as the smooth parameter approaches zero; \( \tilde{D}_{r,t}(\beta;\lambda) \) approaches 1 if \( Health_{r,t} > 0 \) and 0 otherwise.

Then, the modified simulated data sets will be \( \{\tilde{y}_{m,t}^w(\beta;\lambda)\}_{m=1}^N, m = 1, \ldots, M; t = 1, \ldots, T \) smoothed by introducing the function of latent variables, and which then enables the use of the modified simulated data to estimate the descriptive statistical model to obtain the parameter:

\[
\tilde{\Theta}_m(\beta;\lambda) = \arg \max \{L(\gamma(\beta;\lambda); z, \Theta)\}
\]

Denote \( \tilde{\Theta}(\beta;\lambda) = \sum_{m=1}^M \tilde{\Theta}_m(\beta;\lambda) / M \) as the average of estimated parameters, and as sample size approaches a large number and the smooth parameter approaches a small number. According to Keane and Smith(2003), \( \tilde{\Theta}(\beta;\lambda) \) converges to a nonstochastic binding function \( h(\beta) \). So, what is left to do is to get an estimate \( \hat{\beta} \) of the structural parameters in order to make \( \tilde{\Theta}(\beta;\lambda) \) and \( \hat{\Theta} \) as close as possible.

**Step 4: Structural parameter estimation:**
Estimate $\hat{\beta}$ of the structural parameters can be obtained by minimizing a metric function measuring the distance between $\Theta(\beta; \lambda)$ and $\hat{\Theta}$. Following Keane and Smith (2003), the metric function can be used as the likelihood ratio. The first step is to obtain a consistent estimate $\hat{\beta}_1 = \arg \max L(v; z, \Theta(\beta; \lambda))$ of the structural parameters. The number of simulated data sets $M$ is chosen as 1 and $\lambda$ is chosen as 0.05 to ensure the smooth objective function. In the second step, $M$ is chosen as 100 and $\lambda$ is chosen as 0.003 to reduce bias. According to Keane and Smith (2003), $\hat{\beta}_2 = \hat{\beta}_1 - \left(J' L_{\Theta} \left(v; z, \Theta(\hat{\beta}_1)\right) \right)^{-1} J' L_{\Theta} \left(v; z, \Theta(\hat{\beta}_1)\right)$ is a consistent and asymptotic estimate of $\beta$. $L_{\Theta}$ is the Hessian of the descriptive model’s likelihood function. $J'$ is the Jacobian estimate of the binding function $H(\beta_1)$.

**Data:**

The data is from the China Health and Nutrition Survey (CHNS) conducted among nine provinces by the Carolina Population Center and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. It was designed to examine the effects of the health, nutrition, and family planning policies and programs implemented by national and local governments. There are seven waves in total from 1989 to 2006. The survey took place over a 3-day period using a multistage, random cluster process to draw a sample of about 4,400 households with a total of 19,000 individuals in nine provinces that vary substantially in geography, economic development, public resources, and health indicators. My subsample consists of rural and suburban families with complete parental labor supply, characteristic and other related information and with children aged from 0 to 17 for 6 years (1991, 1993, 1997, 2000, 2004, and 2006). The sample size is 6,030, in which there are 3,509 women engaging in agricultural work, 1,858 women engaging in nonagricultural work, 506 women engaging in both types of work, and 557 women not working. Each year’s agricultural production, family income and individual wages are adjusted by the price index of that year. The price index is from the urban rural-specific price index of CHNS using 1998 as the base year.

I find that mothers’ average nonagricultural working time is less than average, agricultural working time in the sample. This is probably because individuals engaging in nonagricultural work are composed of a smaller
proportion of the sample. Mother’s labor income is also less than father’s labor income. Because I use the subsample of rural areas, mothers’ average education is only 6.6 years, in which around 50 percent individuals receive less than 6 years of education. As I consider that it is beneficial to have grandparents at home to supplement maternal care, I also introduced the dummy variable indicating whether a grandparent is nearby. It is found that most elders live close to their grandchildren. From the descriptive statistics of family structure, it is found that there are 55.6 percent single children within the sample, which is roughly in accordance with the one child policy in China. Although the one child policy has been in effect for many years, it is also possible to observe multiple children in the rural areas. Female children are composed of 46.5 percent of the sample, which indicates an approximately even sex ratio.
Reference:


Expenditure Patterns: Evidence from Cote d’Ivoire”, BULLETIN, Vol.57, pp.77-96.


