Who Looks after the Kids? The Effects of Childcare Choice on Early Childhood Development in China*

Jing Zhang,† Simon Appleton,‡,§ Lina Song†,†† and Bing Liu‡‡

†*Business School, University of Nottingham, Nottingham, NG8 1BB, UK (e-mail: j.zhang@nottingham.ac.uk)*
‡*School of Economics, University of Nottingham, Nottingham, NG7 2RD, UK*
§*School of Economics, University of Nottingham Ningbo China, Ningbo, 315100, China (e-mail: simon.appleton@nottingham.edu.cn)*
††*Business School, University of Nottingham Ningbo China, Ningbo, 315100, China (e-mail: lina.song@nottingham.ac.uk)*
‡‡*Aston Business School, Aston University, Birmingham, B4 7ET, UK (email: b.liu7@aston.ac.uk)*

Abstract

This paper examines whether childcare choice affects the early childhood development of children aged 7–59 months. Using the data from Chinese Family Panel Studies, we look at household choices between parental and grandparental cares and the timing of four key early life achievements – walking, talking, counting and toilet training. We conceptualize early childhood development within a household production model, which enables us to identify the impacts of childcare. Our results suggest that compared with parental care, grandparental care delays the achievement of all four outcome measures. Grandparental care is particularly disadvantageous for children who are 'left-behind' by migrant parents.

I. Introduction

Child survival, growth and development are influenced by three underlying factors: nutrients, health and sanitation services, and childcare (UNICEF, 1990; Engle, Menon and Haddad, 1999; Alderman, 2007). While there has been extensive literature examining the impact of the first two factors on child development, research on the effect of childcare

in developing countries is limited. Family's choice of childcare determines child environments which would significantly shape children's abilities and skill formation (Cunha and Heckman, 2007). Scientific research finds that the early years are the most critical period for children's brain development. Various levels of such development are reflected in children's acquisition of physical abilities, social–emotional functionings and language-cognition (Grantham-McGregor *et al.*, 2007, 2014). Early childhood development influences wellbeing throughout life, affecting school attendance, wages, employment, early motherhood and participation in crime, among other things (Irwin, Siddiqi and Hertzman, 2007; Bernal, 2008). The quality of childcare provided in early years is therefore important, with evidence that maternal support in the first few years is strongly associated with child's brain development and mental health (Bowlby, 1951; Luby *et al.*, 2016). This paper looks at how the family's childcare choice affects early childhood development in China.

The importance of maternal care raises an issue for many countries where women's labour market participation grows steadily. Dual-earner families usually seek support from others to care for children while working fulltime. This challenge is more important for parents in developing countries, where there are limited universally accessible and affordable childcare provisions and little regulation for maternity protection (Stumbitz *et al.*, 2018). Evidence from China, the United Kingdom and several other countries shows that the number of children in grandparental care has increased dramatically over the last decade (Goodfellow and Laverty, 2003; Tan *et al.*, 2010; Dong and Zhao, 2017). The contributing factors in China include high female labour force participation (about 70% according to The World Bank, 2018), short average duration and low coverage of maternity leave (Dong and Zhao, 2017), low availability of formal group care provisions (e.g. nurseries) for children under 3 years (Du and Dong, 2013) and unregulated non-group provisions (e.g. nannies). Several studies of China find grandparental care is negatively correlated with child welfare and health outcomes (Ye and Pan, 2011; Mu and de Brauw, 2015; Yue, Sylvia and Bai, 2016). This paper extends these studies by investigating whether there are adverse effects of grandparental care on early childhood development in China. Some studies look at the effects of non-parental, informal care (relatives, friends, non-relatives) as a substitute for mothers' care (Gregg *et al.*, 2005; Bernal and Keane, 2011). Relatively little is known about the impact specifically of grandparental care on early childhood development (see Hansen and Hawkes (2009) and Del Boca, Piazzalunga and Pronzato (2018) for analysis of UK data). To the best of our knowledge, ours is the first study of the impact of informal care (by grandparents) on early childhood development in a developing country.

Using the China Family Panel Studies in 2010, 2012 and 2014, we examine the impact of childcare on the age in months when children achieve four key abilities – to walk, to talk, to count and to use the toilet independently. Each of the four abilities respectively represents the milestone of early development in physical skills, language skills, numerical skills and self-adaptive skills. Given the low availability of formal childcare for the young age group in China, we only distinguish childcare between parents (mainly mothers) and grandparents.

Estimating the impact of childcare on child development is problematic because such childcare decisions are conditioned by many other socioeconomic factors with direct effects on children's early development. With mothers being the traditional primary carer, conventional income and substitution effects apply when considering their labour

supply. Poverty may drive mothers of young children to seek employment while prosperous mothers might enjoy maternity leave or become homemakers. Conversely, women who are better educated or from an advantaged background may have superior opportunities in the labour market and so delegate daytime childcare responsibilities to others. While it is possible to control for some observable factors affecting labour supply and childcare choice, unobserved factors may remain. To establish the causality of childcare choice on the ages in months at which the skills are obtained, we model each skill using two-stage instrumental variable estimations in a semi-parametric proportional hazards model (Terza, Basu and Rathouz, 2008).

The paper proceeds as follows. In section II, we firstly discuss the context and issues related to the measurements of child outcomes and childcare in China. Section III presents a conceptual framework and estimation strategy before section IV, which describes our data. Section V reports our main results, including specifically on 'left-behind' children, while section VI concludes.[1]

II. Context and issues

Measurements of early childhood development

We measure childhood development by the recorded month in which children first achieve the ability to walk, to talk, to count and to use the toilet independently.

Walking independently by 24 months of age is one of six distinct gross motor milestones identified by the World Health Organization (WHO Multicentre Growth Reference Study Group, 2006). Recent research finds that early motor skills are important not just for later motor skills but also for adaptive and cognitive development (Ghassabian *et al.*, 2016). Oudgenoeg-Paz, Volmsagean and Leseman (2012) suggest that early acquisition of walking at the age between 12 and 20 months propels infants' language development during 16–24 months, and Walle and Campos (2014) also confirm that acquisition of walking during 10–13.5 months of age is associated with a significant increase in receptive and productive language. Independent walkers usually produce more social referencing and joint attention, as well as increased vocalizations, directed gestures and social interaction bids (Clearfield, 2011; Leonard *et al.*, 2015). Walking is also linked to emotional changes as walking infants show increased elation, together with greater levels of wilfulness (Biringen *et al.*, 1995).

Children's language development in vocabulary and word combinations is another important achievement of early childhood. In this paper, we measure the age of talking using the month in which the child can form a short but meaningful sentence, such as 'I want to eat (*wǒ yào chīfàn*, expressed using four Chinese characters)'. Children who have limited expressive vocabulary (usually <50 words) and have difficulties in combining two to three words at 24 months of age are identified as late talkers (Rescorla, 1989). Del Boca *et al.* (2018) find that children cared by grandparents at the age of 18 months have better vocabulary skills like naming objectives but this effect turns to be insignificant after controlling for the endogeneity of childcare. In our research, we test whether grandparents have similar effects on children's language development.

---

[1]An online appendix provides further material with details about: descriptive statistics; first stage results (instrumentation of childcare choice); disaggregations by gender and rural–urban location; and robustness analysis.

Counting from one to 10 is the third measurement of child ability. Numeracy is particularly valued by parents in China (Aunio *et al.*, 2008). The benefits of early year's mathematics learning have been advocated by several researchers: the National Research Council (2009) recommends that children should start to learn math as early as three years old.

Our final indicator of child development – toilet training – is regarded as a milestone of children's adaptive (self-help) development. In this study, being successfully toilet trained means that the child is able to independently complete the process of taking off pants, urinating and putting on pants. The completion of toilet training marks a transition from the early infanthood stage to becoming more independent. In the 20th century, prescriptions for toilet training swung like a pendulum between the polar opposites of passive permissiveness and systematic control. However, the current consensus is that professionals and childcarers should be concerned if a child refuses to be toilet trained or the achievement is delayed, behavioural management of toilet training being recommended to reduce the adverse implications in such cases (Luxem and Christophersen, 1994).

Childcare mechanisms

Early childcare is usually provided by three mechanisms: parents (assumed to be primarily the mother), relatives (typically the grandparents), and formal childcare providers including group-based (nurseries) and non-group-based (childminders, nannies). In this study, we only consider parental and grandparental cares because institutional childcare is not commonly available for children under 3 in China and the sample size of children in formal childcare is too small to be included in the empirical analysis.

Studies in neurology and psychology examine the impact of maternal support in the first few years (Luby *et al.*, 2016) and infant-mother attachment (Bowlby, 1951), both of which show strong associations between children's brain development and their long-term behavioural, emotional and cognitive development. It is widely believed that parenting, particularly mothering, is a strong and consistent predictor of child outcomes (The NICHD Early Child Care Research Network, 1998; Belsky *et al.*, 2007). However, young mothers face a choice between caring for children and paid employment. Thus, non-maternal childcare mechanisms are established to substitute for maternal care.

In developed countries, formal childcare providers such as nurseries, qualified childminders and nannies are important early childcare mechanisms. There is evidence that high-quality childcare not only neutralizes the potential negative effects of maternal employment but even produces positive effects on children's cognitive development (Felfe, Nollenberger and Rodríguez-Planas, 2015). However, these childcare mechanisms are less common and weakly regulated in developing countries such as China. China started childcare reform in the 1990s, mainly in the form of kindergarten/nursery privatization, which resulted in a dramatic decline of public-funded nurseries (Du and Dong, 2013). Public kindergartens are mostly commercialized, and the number of private kindergartens rose significantly, with diverse quality of services. In addition, national regulations only allow kindergartens to recruit children aged 3 years and above. Increasing numbers of urban families rely on domestic workers/nannies to care for children under three. However, the proportion of families using domestic workers is still small, and there are no professional

standards or regulations.[2] In our data set, there were too few cases of formal childcare to warrant, including it as a category in our analysis.[3] Thus, it is grandparents and other relatives who typically provide childcare when the mothers are working. Dong and Zhao (2017) report that in urban China, the proportion of children aged under seven in nurseries and other institutions dropped from 26% in 1991 to 22% in 2011, while the percentage cared for by grandparents increased from 40% to 56%. Care by grandparents is also common in rural China and has risen even faster than in urban areas due to mass rural–urban migration. The phenomenon of 'left-behind children' is a major factor causing the proportion of rural children aged 0–6 cared by grandparents to increase from 27% in 1988 to 49% in 2011.

China's maternity leave policy is another reason for grandparents being important carers of preschool children. The duration of paid maternity leave was 90 days before 2012 and then extended to 98 days, which is in line with the standard of the International Labour Organization. Many provinces granted extra days of late childbirth leave to women who had the first childbirth at the age of 24 or older, which makes the total duration of maternity leave vary across provinces from 90 days to 180 days (Jia, Dong and Song, 2018).[4] China's maternity leave is childbirth related and solely for mothers. It is different from and shorter than the policy in many OECD countries that includes extra childcare-related leave, which, in some countries, can be taken by either parent. Besides, maternity leave provision does not benefit all employed mothers (Dong and Zhao, 2017). Employees with a short-term contract in the state sector do not have such benefits, and the coverage of paid leave in the private sector is much lower as a result of no effective means of enforcing the regulations. Jia *et al.* (2018) find that during 1998–2008 only about 60% employed urban mothers took paid maternity leave. For employed mothers with a college education, 88% took the leave, and the average duration was 107 days. In contrast, among employed mothers without a college education, only 36% took the leave, and the mean duration for these mothers was just 39 days. It is estimated that only 33.5% of female employees in cities were covered by maternity insurance in 2012 (Dong and Zhao, 2017). Women who were not covered by maternity insurance were more likely to have lower socioeconomic status in the informal or private sectors, where paid maternity leave is generally unavailable.

The pattern of early retirement in China – often around age 55 for women – also encourages the reliance on grandparents for childcare. In the dynamic economy of China, younger women recently educated may be more productive in the labour market than their aging parents, who received less education. Consequently, grandparenting becomes an important substitute for mothers' care in many Chinese families.

[2]In June 2019, China's State Council issued a Guidance for Promoting the Quantity and Quality of Domestic Services, which is the first policy in domestic service sector and mainly a blueprint of future development.

[3]In the questionnaire, there is a direct question to the respondent about the primary daytime carer, including six categories: mother, father, paternal grandparents, maternal grandparents, nursery/kindergarten and domestic worker. We group maternal and paternal care into parental care category (the proportion of paternal care is very small). We group paternal- and maternal-grandparental care as grandparental care. The proportion of children cared by kindergarten and domestic workers is too small (less than 0.4% of each sample size). Therefore, we only consider parental and grandparental care in the empirical analysis.

[4]From 2016, China's one-child policy that lasted more than three decades was eliminated. To encourage the childbirth, many provinces adjusted the maternity leave policy to '98 days national standard leave + child birth bonus days'. Nowadays, the duration of paid maternity leave is between 128 days and 190 days. However, year of 2016 is out of our observed period in this study that is from 2010 to 2014.

Tied by close family bonds, grandparents in China typically offer childcare free of charge to the parents. However, there are concerns over the quality of care provided, as grandparents are usually less educated and hence less likely to adopt modern childcare practices than other carers. Hansen and Hawkes (2009) found that children in the United Kingdom with grandparental care usually do better in vocabulary skills although Del Boca *et al.* (2018) analysed the same dataset and overturned this result after controlling for the endogeniety of childcare choice. Both studies report children cared for by grandparents are weaker in other cognitive development and have more behavioural problems. Compared to parents, grandparents might also have lower awareness of healthy means of raising up children. Li, Adab and Cheng (2015) find that children who live with two or more grandparents have a 70% higher risk of being overweight or obese than those who do not. Similarly, Mo *et al.* (2016) find a higher prevalence of being overweight or obese among left-behind preschool-aged children but also find the risk of wasting (being too thin) is higher among that population. In this study, we examine, comparing to parental care, whether grandparental care has similar effects on the development of the four milestones in the first few years of children's life.

III.   Conceptual framework and estimation strategy

Conceptual framework

In this section, we adapt Becker's household production model to child-rearing, providing a simple representation of some of the key issues involved in modelling early childhood development. In this approach, early childhood development is modelled as being produced by the household using goods and time inputs (Ben-Porath, 1967; Leibowitz, 1974). We start with a child development production function, equation (1), where the early development achievements, A, of child i are determined by choices over inputs made by parents and by control variables:

$$A_i = A_i(C_i, TP_i, TG_i, E, X, v) \quad i = 1, \ldots, N. \tag{1}$$

For exposition, we assume that all spending on children, C_i, is input into child development. This would include food, books, toys and health care (we abstract from child expenditure that has no development benefits). With time inputs, T, to childcare, we distinguish the time provided by two types of caregivers, parents (TP) and grandparents (TG). We do not consider other childcare mechanisms here but this model can be easily extended to include more categories of childcare. The control variables in equation (1) are parental education (E) and other observed parent and child characteristics (X). There are also unobserved determinants of child achievement, v, which comprise the characteristics of the child, parents and environment that are not measured in the survey.

We assume the household chooses the amount of good and time inputs to allocated to child development based on maximizing a household utility function subject to constraints. For example, a simple utility function would be:

$$U = U(C_p, C_i, L, A_i), \tag{2}$$

which includes the child's abilities (A_i), the consumption of goods and services (C) by parents, p, and the child i, and leisure (L). Equation (2) assumes a nuclear household consisting of parents and one child. It could be generalized to several children, who may have different weights in the household utility function (e.g. if there is pro-son bias). For simplicity, we assume a unitary household, abstracting from potential differences between parents.

The household maximizes utility with respect to consumption and time subject to child development production function (equation (1)), a time constraint for parents (equation (3)), and a budget constraint (equation (4)):

$$TT = TW + TP + L \tag{3}$$

$$C_p + C_i + P_{TG} \cdot TG = W \cdot TW + M \tag{4}$$

TT is the total time available to parents, allocated to wage work (TW), childcare (TP) and leisure (L). For simplicity, we use a one-period model, where spending equals income. Household expenditure comprises the general consumption of parents and children together with the cost of childcare provided by grandparents, at a price of P_{TG} per unit of time. As note previously, grandparents' time typically is not charged for at a monetary rate – however, we use P_{TG} to capture the opportunity cost.[5] Household income is the sum of the wages (W) paid to parents and unearned income (M). The wage (W) for parents depends on their education level and other personal characteristics.

$$W = W(E, X) \tag{5}$$

For a household, parental care and grandparental care are imperfect substitutes. How much of each kind of carer's time is allocated will depend on their relative productivity in child development ($\delta A / \delta T$) and the relative cost of their time. At optimum, the marginal rate of technical substitution between the two caregivers will equal the ratio of the cost of their time.[6] *Ceteris paribus*, parents who could earn high wages would be more likely to rely on their grandparents to look after their children.

The solution to the optimization problem can be represented by reduced form demand equations whereby all the choice variables – consumption, leisure and time allocated to child-rearing – are a function of all the exogenous variables:

$$C_p, C_i, L, TP_i, TG_i = f(P_{TG}, E, X, M, v) \tag{6}$$

Attempting to estimate directly the structural relation – the child development production function – in equation (1) is problematic due to the presence of endogenous inputs. The presence of unobservable v in both equations (1) and (6) implies that ordinary least squares (OLS) would suffer from endogeneity bias. One solution to this is to use an approach with instrumental variables (IV).

---

[5] In a more realistic model, one would allow for intra-family bargaining and a time constraint on grandparents.

[6] To derive the optimal time and goods allocation to childcare, equations (3) and (4) can be combined into a single 'full income' constraint, and a Lagrangean function formed from the objective function, equation (2) and this constraint. Differentiating Lagrangean with respect to the choice variables, we obtain the following tangency condition for optimum time allocation to childcare: $(\partial A / \partial TP) / (\partial A / \partial TG) = W / P_{TG}$.

Estimation strategy

To address the endogeneity of childcare choice, we employ a strategy of two-stage esti-
mation with IVs. However, the standard two-stage least squares (2SLS) estimation is not
suitable for our data due to the nature of limited dependent variables in the first and second
stages. In our data, we do not have continuous variables to measure the childcare time input
(*TP* and *TG*). Instead, childcare is recorded as categorical information on the type of the
main daytime carer of the child. This means that our childcare choice variable is binary,
taking a value equal to 1 if a child is cared by a grandparent, and 0 by a parent. Thus, in
the first stage, it is appropriate to model childcare using a binary choice model such as a
probit model. In the second stage, our measures of early childhood development are the
ages at which children reached four milestone achievements (walking, talking, counting
and independent toileting). Each achievement is an event whose timing is to be modelled.
The four timing variables are right censored: in our sample, some young children may
not have attained a specific achievement but will do so later. Hence, the appropriate way
to model equation (1) is using survival analysis which is commonly used for modelling
duration data, in the case, the time until a child attains a specific achievement. Therefore,
childcare choice and child outcome are both measured by limited variables, and both first
and second stages are nonlinear models.

    Under this circumstance, we adopt two IV-based approaches to correct for endogeneity
bias in nonlinear models – two-stage predictor substitution (2SPS) and two-stage residual
inclusion (2SRI) (Terza *et al.*, 2008; Tchetgen Tchetgen *et al.*, 2015). The 2SPS estimator
is a straightforward analogous to 2SLS, using the predicted values computed from the first
stage regression (i.e. predicted probabilities of parental care and grandparental care in this
paper) to substitute the endogenous variable (i.e. the dummy of childcare choice) in the sec-
ond stage. The 2SRI estimator has the same first stage regression as 2SPS; however, in the
second stage, the endogenous variable is not replaced. Instead, the first stage residuals are
used as an additional control in the second stage estimation. The 2SRI method, also known
as the control function approach, is essentially the same as Durbin–Wu–Hausman endo-
geneity test in the linear context (Hausman, 1978; Wooldridge, 2014). Although the 2SPS
has been widely used in health economics and epidemiology research, Terza *et al.* (2008)
find that 2SRI method provides consistent estimates while 2SPS does not. In this paper,
we adopt both methods but use the results from the 2SRI estimator as the main findings.

    Aside from the central concern over childcare choice, we also investigate the endogene-
ity of two other determinants of childhood achievement – parental education and household
income. Although parental education was assumed to be fixed in the simple conceptual
model presented previously, the observed effects of education may reflect pre-existing abil-
ity or parental background rather than schooling *per se*. We use household income, I as a
proxy for the consumption of child goods that are inputs to child development, C_i, as we
do not have direct measures of spending on child goods. Household income is potentially
endogenous because it depends on childcare choice and work decisions, as well as the wage
level, which might be affected by unobserved parental ability endowment that is correlated
with the child outcome.

    To deal with the endogeneity of childcare, parental education and household income,
we use a set of instrumental variables that are correlated with the endogenous variables but
have no influence on the child development measurement. Our first IV proxies the local

demand condition that influences childcare decisions and household income. Similar to Bernal and Keane (2011), we use the 20th percentile annual wage of female employees in the household's county (IV1). The local female wage rate will influence the opportunity cost of parental childcare, which is typically provided by mothers. The lower percentile of the wage rate is used to capture the wage paid to unskilled labour, because we do not want the variable to capture the return on human capital investment, as this could affect the benefit to investing in early childhood development as well as the cost. The second is a dummy variable whether the parent was born before or after 1978 (IV2) to capture a policy change on school-aged children since 1985, the year when China introduced 9-year compulsory education system, re-structured the secondary school education and increased autonomy of higher education institutions (Cheng, 1986). On average, parents who were born after 1978 are more likely to have a higher education level and a higher opportunity cost of caring the child themselves, and hence prefer to ask grandparents to fill the childcare gap. The next instrument is the education of the grandparents (IV3) which would influence parental education and through education to income.[7] Our final IV is an index for household assets (IV4), generated by principal component analysis, which reflects the possession of motor vehicles, house and second house, as well as jewellery. Household assets are correlated with household income but are unlikely to be affected by the short-term labour supply decisions modelled in our conceptual framework.

To test the validity of the IVs, we calculate the covariance between each instrumental variable (Z) and the residuals, v obtained from naïve estimation of equation (1), that is $Cov(v, Z)$. For all the instruments, the covariances are close to zero, indicating instruments exogeneity (Wooldridge, 2014). Besides, we estimate our models by including three of the four IVs each time to test whether the estimated coefficients are consistent (Wooldridge, 2014).[8]

The three equations in the first stage are hence:

$$\Pr(G = 1 | X, Z) = \Pr(X\alpha_1 + Z\alpha_2 + \varepsilon_R > 0) \tag{7}$$

$$E = X\beta_1 + Z\beta_2 + \varepsilon_E \tag{8}$$

$$\ln I = X\theta_1 + Z\theta_2 + \varepsilon_I \tag{9}$$

respectively, where G is the childcare choice dummy taking a value of 1 if a child is cared by a grandparent and 0 by a parent; E refers to parental education, I is household income, and Z is a set of instrumental variables as discussed above. X represents observed parent/child characteristics, such as age, child sex, birth weight, whether the child is the first born, only child dummy, residential status and regional dummies.[9] ε_R, ε_E and ε_I are respectively, the

---

[7]A priori, the exclusion restriction required for this instrument is questionable: the education of grandparents may affect children in ways other than through parental education, particularly if the grandparent is one of the child's carers. However, preliminary estimates found grandparental education was not a significant determinant of child achievement in our data. The validity test of IVs (in the online appendix) also confirms this.

[8]This is similar to a test of overidentification restrictions as stated in Wooldridge (2014) p.428–429). The results are reported in the online appendix.

[9]Birth weight could be regarded as endogenous, reflecting parental health and nutrition. However, we include it to control for genetic endowments and longer term influences on child development. Parental education may also be endogenous from a life cycle perspective and we address this empirically.

residuals in each regression. Equation (7) is estimated using a probit model while equations ((8)) and (9) are estimated using OLS. We then calculate the predicted probabilities of a grandparent being the main carers of the child ($\hat{G}$), the predicted parental education ($\hat{E}$) and the predicted household income ($\hat{I}$). We also calculate the three residuals for each child $\hat{\varepsilon}_R$, $\hat{\varepsilon}_E$ and $\hat{\varepsilon}_I$, each defined as the difference between the observed and predicted values of the endogenous variable.

In the second stage, we model the timing at which each of the four abilities is obtained. Following Cox (1972), we use a semi-parametric proportional hazards model, where the hazard rate – the probability of attaining ability A at time t given that it has not previously occurred – for a child i is given by:

$$\lambda_i(t|\mathbf{Y}) = \lambda_0(t)\exp(\mathbf{Y}\gamma) \tag{10}$$

where $\mathbf{Y}$ is the vector of determinants of childhood development, and $\lambda_0(t)$ is a baseline hazard. An advantage of the Cox approach is that it does not impose a distribution on the baseline hazards.

The log of the hazard function is a linear combination of parameters and repressors, that is

$$\log\lambda_i(t|\mathbf{Y}) = \alpha_0(t) + \mathbf{Y}\gamma \tag{11}$$

When using the 2SPS method, the empirical model in the second stage is:

$$\log\lambda_i(t|\mathbf{Y}) = \gamma_0(t) + \gamma_1\hat{G} + \gamma_2ln\hat{I} + \gamma_3\hat{E} + \gamma_4X + v_{2SPS} \tag{12}$$

whereas the empirical model of the 2SRI method is:

$$\log\lambda_i(t|\mathbf{Y}) = \rho_0(t) + \rho_1 G + \rho_2 I + \rho_3 E + \rho_4 X + \rho_5\hat{\varepsilon}_G + \rho_6\hat{\varepsilon}_E + \rho_7\hat{\varepsilon}_I + v_{2SRI} \tag{13}$$

The results of 2SRI is analogous to Durbin–Wu–Hausman test so can be used as a test of endogeneity of our key regressors – the dummy of childcare, parental education and household income. For all regressors, we reject the null hypothesis of exogeneity, that is ρ_5, ρ_6 and ρ_7 are statistically different from zero in most of the specifications.

We estimate equations (7)–(9), (12) and (13) separately using Stata. Standard errors obtained may fail to appropriately account for extra variability due to the first stage estimations and might understate uncertainty. To produce more accurate estimates of standard errors, we perform all the estimations based on 500 bootstrap samples.

IV.   Data and variables

The data we use are from the China Family Panel Studies (CFPS) of 2010, 2012 and 2014, nationally representative longitudinal social surveys conducted by the Institute of Social Science Survey at Peking University, China (Xie and Hu, 2014). It includes a section on child development, asking about the timing of the four milestone achievements as described in previous sections. We collect individual and household data for children whose child outcomes were reported by the parent or the main carer. For each child, we take one observation when modelling each achievement. Since the four achievements are developed at different age ranges, we select children who are aged 7–24 months for

TABLE 1

Descriptive statistics on measurements of early childhood development and childcare

	Walking	Talking	Counting	Toilet training
Child age range in month	7-24	12-36	18-59	18-59
Whether the ability has been obtained (1= Yes; 0 = No)				
Mean	0.399	0.217	0.172	0.271
S.D.	0.490	0.412	0.377	0.445
Age of obtaining the ability (month)				
Mean	13.346	21.322	30.483	29.283
S.D.	2.291	5.231	8.042	7.577
First Quartile	12	18	24	24
Third Quartile	14	24	36	36
Childcare choice (%)				
Parents	**66.94**	**57.16**	**53.39**	**51.71**
by mothers	65.76	55.77	51.71	49.8
Grandparents	**33.06**	**42.84**	**46.61**	**48.29**
at least 1 parent at home	27.73	33.58	32.36	35.43
no parents at home	5.33	9.26	14.25	12.86
Observations	1107	1361	1845	1524

Notes: Statistics are weighted. Among the children who are cared by grandparents, if the child lived with either parent more than 5 months over the past 1 year of the survey, we define the child as living with at least one parent at home; otherwise, the child is living with no parent at home (Zhou *et al.*, 2018).
Sources: China Family Panel Studies, 2010, 2012, and 2014

walking, 12–36 months for talking, and 18–59 months for counting and toilet training on the date of the survey. To ensure that the childcare choice is measured at the time when the ability is developing, we only keep the children who are developing a particular ability at the time of the survey (i.e. the ability is not yet obtained), and the children who have obtained the ability within 3 months of the survey. This limits retrospective measurement errors with the outcome variables.[10] We exclude children whose parent did not report their childcare pattern or whose records were obviously erroneous, for example, being able to walk independently before 6 months or talk in a full sentence before 12 months. Since a child might obtain one ability in one wave of the survey and other abilities in the next waves, we construct separate samples for each ability. Definitions of variables and the descriptive statistics are available in the online appendix.

Table 1 shows the typical ages at which children show the capability to 1) walk alone at least three to five steps without any assistance; 2) form a short meaningful sentence; 3) count from 1 to 10; and 4) be successfully toilet trained. The age at which children first walk shows the least variation: the mean is 13 months within an inter-quartile range of 2 months. The ages at which the other three achievements are reached show more variation: the coefficient of variation (the standard deviation divided by the mean) is 0.17 for walking

[10] For a robustness check, we expanded the sample by including children who have achieved the ability within 6 months of the survey, assuming that the main caregiver is not changed in the past half of a year. See the results in the online appendix.

but 0.25–0.26 for the other three abilities. The 21st month is the mean age at which children talk in full sentences, but the first quartile is 18 months and the third quartile 24. Counting to 10 and independent toileting are achieved at similar ages, the means being the 30th and 29th months respectively.

Table 1 also reports the type of the main daytime carer. The breakdowns across our two categories vary according to the child's age, with mothers being the most common at younger ages (about two third in the walking sample) and grandparental care becoming concentrated in later years. Regardless of the daytime carer, the majority of children are living with at least one parent (mostly including mothers). The proportion of children who are left behind by both parents increases from about 5% in the walking sample to about 13–14% in the counting and toilet training samples. Overall, the majority of children (more than 50%) are primarily cared for by their parents (almost all by mothers) but grandparents become equally important carers as mothers for children above 18 months.

V.   Results

Main results

Table 2 presents the main results from the Cox models for the hazards of obtaining the four abilities of children (equation (11)). For each ability, we report the results from the naïve estimates without tackling the endogeneity of relevant regressors, the estimates using IVs from 2SRI method (equation (13)) and 2SPS method (equation (12)). Our estimates from 2SRI and 2SPS are similar but largely differ from the naïve estimates. The results of Durbin–Wu–Hausman test suggest that childcare, mothers' education and household income are endogenous regressors so that the naïve estimates are biased. The table also reports that the covariances between the residuals from the naïve estimates and each instrumental variable, that is, $Cov(res_{-n}, IV)$, are close to zero, indicating the exogeneity of all the IVs.

We use the estimates from 2SRI as our main results. In general, being cared for by grandparents has a negative effect on the hazard for all the abilities – walking, talking, counting and being toilet trained. Since the survival model is a proportional hazard one, the coefficients can be interpreted as scaling the baseline hazard of a particular child development achievement.[11] Overall, the results suggest that parental care, other things being equal, is better for early childhood development than being cared for by grandparents. These findings are broadly consistent with previous research about the adverse effect of grandparental care, with the possible exception of talking.[12]

These findings are only evident after controlling for endogeneity. Naïve estimates find the only negative effect of using grandparents is on counting. The differences between the naïve and IV estimates imply a positive association between the unobservables associated with grandparental care and those associated with early childhood development. Intuitively, parents who have unobserved advantages in terms of their work productivity – leading them to use grandparents – may also tend to have unobserved advantages when it comes to child

---

[11] For example, the coefficient of grandparental care (-3.132) in column (2) implies that such care cuts the hazard by 4.4% ($0.044 = exp(-3.132)$) relative to parental care and therefore a longer time to obtain the ability, *ceteris paribus*.

[12] Hansen and Hawkes (2009) find that, in the United Kingdom, grandparental care benefits child vocabulary skills. However, Del Boca *et al.* (2018) find this result disappears after controlling for the endogeneity of care.

TABLE 2

Main results: survival models for timing of early childhood achievements

	Walking			Talking			Counting			Toilet training		
	Naïve (1)	2SRI (2)	2SPS (3)	Naïve (4)	2SRI (5)	2SPS (6)	Naïve (7)	2SRI (8)	2SPS (9)	Naïve (10)	2SRI (11)	2SPS (12)
Grandparental care	−0.006	−3.132***	−3.133***	0.081	−2.870***	−2.870***	−0.235**	−5.661***	−5.757***	−0.141	−4.056***	−4.044***
	(0.091)	(0.472)	(0.480)	(0.115)	(0.794)	(0.832)	(0.117)	(0.980)	(0.904)	(0.094)	(0.672)	(0.633)
Mothers' education	0.023*	0.332***	0.335***	0.029*	0.261***	0.256***	0.057***	0.089	0.092	0.011	0.223***	0.224***
	(0.012)	(0.059)	(0.061)	(0.017)	(0.058)	(0.063)	(0.018)	(0.064)	(0.060)	(0.014)	(0.054)	(0.048)
Log(Household income per capita)	−0.001	−0.956***	−0.971***	0.023	−0.079	−0.053	0.055	0.662*	0.618*	−0.031	−0.379	−0.384
	(0.037)	(0.261)	(0.258)	(0.050)	(0.274)	(0.276)	(0.054)	(0.345)	(0.335)	(0.038)	(0.300)	(0.277)
1st-stage residuals		Yes			Yes			Yes			Yes	
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1107	1107	1107	1361	1361	1361	1845	1845	1845	1524	1524	1524
Log likelihood	−2632.87	−2541.82	−2543.46	−1822.51	−1768.61	−1769.61	−1928.19	−1882.78	−1886.97	−2512.79	−2439.08	−2439.19
Cov(res_n, IV1)†	−0.023			−0.011			0.015			0.006		
Cov(res_n, IV2)	0.000			0.002			−0.001			0.000		
Cov(res_n, IV3)	0.114			0.005			0.002			−0.022		
Cov(res_n, IV4)	−0.009			0.015			0.003			0.007		
Durbin–Wu–Hausman (P-value)‡												
$\rho_5 = 0$ (for $\hat{\varepsilon}_G$)		0.000			0.000			0.000			0.000	
$\rho_6 = 0$ (for $\hat{\varepsilon}_E$)		0.000			0.000			0.597			0.000	
$\rho_7 = 0$ (for $\hat{\varepsilon}_I$)		0.000			0.711			0.050			0.200	

Notes: Standard errors with 500 bootstrap replications are reported in parentheses. Original endogenous variables and first-stage residuals are used in two-stage residual inclusion (2SRI) method. Predicted endogenous variables are used in two-stage predictor substitution (2SPS) method. Control variables include mother's age, child gender, birth weight, a dummy of whether the child is the first born, a dummy of whether the child is the only child in the family, the dummy variables of household resident status ('rural resident' and 'rural to urban migrant' with 'urban resident' as the base group), and regional dummy variables ('central provinces' and 'western provinces' with 'coastal province' as the base group.

†Covariance between the residuals from naïve regression (res_n) and each instrumental variable is calculated. IV1= 20 percentile wage of female employees in the county; IV2 = dummy of whether the mother is born after 1978; IV3= maternal grandparents' education, and IV4=household assets index. All the covariances are close to zero, indicating the exogeneity of IVs.

‡ The results of Durbin–Wu–Hausman test confirm grandparental care, mothers' education and household income per capita are endogenous.

*P < 0.1, **P <0.05, ***P <0.01

development. Not controlling for the endogeneity of the childcare choice, therefore, makes it appear that parental care is less beneficial than it actually is.

Turning to the other two endogenous regressors, mothers' education has a positive and significant effect on the hazards of walking, talking and being toilet trained.[13] Each year of education raises the baseline hazard by 24–40%. The estimated effect on the hazard for counting is positive but not significant in the IV estimates.

Household income delays the development of walking skill but hastens counting skill. It does not affect talking and toilet training. One possible explanation of the negative effect on walking is that, in the walking sample, children in high-income families (above the median in per capita income) are on average about 500 grams heavier than those in the low-income families (below the median) and heavier children often walk later.[14] Among the four selected abilities, counting from 1 to 10 may be regarded as the most obvious result of deliberate educational training. The income effect is large: doubling household income increases the hazard of learning to count by 94%.[15] Our data also show that rich families tend to invest more time in teaching numbers to children aged between 1 and 3 years.[16] This may foreshadow subsequent increases in human capital investment and be positive overall for China, given its rapid economic growth. However, it may also be one channel for the intergenerational transmission of inequality within the country.

Subgroup analysis

Our main results show strong negative effects of grandparental care on all the early development outcomes. However, estimating the average effect of childcare choice may conceal the differences across subgroups. For example, Magnuson *et al.* (2016) suggest that early childhood education programmes have a roughly equal impact for boys and girls on cognitive and achievement measures but boys gain more on other school outcomes. Del Boca *et al.* (2018) find the negative effect of grandparental care is only significant for children from a more disadvantaged background. To assess how the effects of explanatory variables vary with different types of child and different types of household, we estimate equations ((12)) and (13) separately for several subgroups, divided by the children's gender and residential status. A bootstrap and permutation test (Cleary, 1999) is used to test whether estimated coefficients on grandparental care of the subgroups are statistically different from each other.[17] For brevity, we only report the coefficients of grandparental care from 2SRI estimates.[18]

---

[13] We use maternal education as a proxy for parental education to simplify the empirical analysis. Results for controlling paternal education are reported in the online appendix.

[14] We also calculated the weight-for-age z score for children and the difference between high- and low-income countries is 0.338 and statistically significant from the two-sample test.

[15] While the coefficient β in the proportional hazards model is an approximation to the proportional effect, a more precise estimate is $exp(\beta) - 1$.

[16] Some families reported how often a family member teaches numbers to the child aged 1–3 years with answers precoded into six bands. The average household net incomes per capita by the frequency bands are: RMB 11,670 (every day), RMB 9,453 (several times a week), RMB 10,116 (several times a month), RMB 6,902 (once a month), RMB 7,368 (several times a year), and RMB 4,614 (never) respectively.

[17] Cleary (1999) method is a bootstrapping procedure to calculate empirical *P*-values that estimate the likelihood of obtaining the observed differences in coefficient estimates with the null hypothesis that the true coefficients are equal for each variable.

[18] Results from first-stage regressions and from 2SPS method are available upon request.

Our main findings confirm the negative effects of grandparental care on all the four abilities for both boys and girls (Panel A in Table 3 refers). The magnitude of the co-efficients implies boys are more negatively affected by grandparental care for walking while girls are more negatively affected for talking, counting and toilet training. However, the Cleary test does not reject the equality of coefficients across the two genders for any ability.

When dividing the sample by rural–urban location,[19] there are heterogeneous effects between the rural and urban samples for the effects of grandparental care (Panel B in Table 3 refers). Grandparental care has a stronger negative effect on the hazard of walking for rural children. Similarly, its adverse effect on language development is significant only in the rural sample, with insignificant effects on urban children. A Cleary test rejects equality of coefficients for both abilities. Interestingly, grandparental care has a stronger negative effect on counting and toilet training for urban children, but the Cleary test suggests the difference is only significant for toilet training. The difference in toilet training may be due to relevant caring practices. Rural children are more likely to be cared using traditional ways, for example, using cloth diapers that need to be washed by hands, so carers (both parents and grandparents) prefer to train the children as early as possible. Therefore, compared to parental care, the average negative effect of grandparental care on toilet training is smaller for rural than urban children. Overall, it appears that grandparental care has negative effects on children living in both rural and urban China; however, it is more of a concern in rural areas in terms of walking and talking.

Heterogeneity in the effects of left-behind children

Considering the phenomenon of mass rural–urban migration and millions of 'left-behind children' in the rural area (Chen *et al.*, 2009; Lu, 2013; Dong and Zhao, 2017), we examine whether left-behind children, particularly those left by both parents, are in a more disadvantaged position. Following the method in Zhou *et al.* (2018), if a child has lived with either parent for at least 5 months out of the previous year, we define the child as living with at least one parent at home; otherwise, the child is living with no parent at home. Then we decompose our childcare choice variable into three categories: parental care (base group same as before); grandparental care with at least one parent living at home (*G1*); and grandparental care with no parents living at home (*G2*).[20] Then we modify the childcare model (equation (7)) in the first stage to a multinomial logit model, including an additional IV – the average proportion of children who are aged 0–59 months and left behind by mothers in the county of residence (IV5). Then the second-stage equation for the 2SRI method becomes

---

[19] Children who are rural to urban migrants are included in the urban model and a dummy variable of migrant is included.

[20] There is no direct question about whether the child is 'left-behind' by one parent or both parents, so we use whether the child lives with parents as a proxy. Our data allow us to identify 'father-only left-behind children', 'mother-only left-behind children' and 'both parents left-behind children'. However, the observations of 'mother-only left-behind' is very small. Since living with mothers is more important for young children in our samples, and more categories will raise more endogenous problems, we only consider two categories for children who are cared by grandparents, that is, living with at least one parent, and living with no parents.

TABLE 3

Subsample results by gender and living place – 2SRI estimates

Panel A (by gender)	Walking		Talking		Counting		Toilet training	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Grandparental Care	-3.256***	-2.852***	-3.236**	-3.583***	-4.080***	-5.969***	-2.095**	-3.678***
	(0.682)	(0.685)	(1.374)	(1.012)	(1.016)	(1.398)	(0.837)	(1.012)
Observations	602	505	761	600	1021	824	835	689
Log Likelihood	-1235.536	-998.732	-828.032	-729.737	-833.008	-814.726	-1115.109	-1032.241
Cleary Test (*P*-value)†	0.296		0.432		0.102		0.106	
Panel B (by living place)	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Grandparental care	-2.257***	-3.903***	-0.972	-4.509***	-6.226***	-5.194***	-4.717***	-1.842**
	(0.806)	(0.601)	(1.206)	(1.292)	(1.474)	(1.338)	(1.017)	(0.790)
Observations	497	610	601	760	719	1126	624	883
Log Likelihood	-1009.721	-1223.860	-635.044	-917.404	-756.068	-905.074	-857.393	-1279.900
Cleary Test (*P*-value)†	0.018		0.006		0.224		0.006	

Notes: Standard errors with 500 bootstrap replications are reported in parentheses. Original endogenous variables and first-stage residuals are used in two-stage residual inclusion (2SRI) method. Results from the first stage and two-stage predictor substitution (2SPS) method are available upon request. Control variables are included.
†Cleary test (Cleary, 1999) is a bootstrap and permutation test with a null hypothesis that estimated coefficients of the subgroups are equal for each variable. *P*-value < 0.05 indicates that the coefficient on Grandparental Care (ρ_1 in equation (13)) differs statistically across subgroups at 5% significance level.
*$P < 0.1$, **$P < 0.05$, ***$P < 0.01$.

TABLE 4

Heterogeneity in the effects of left-behind children – 2SRI estimates

		Walking	Talking	Counting	Toilet training
All - naïve	Grandparents_at least one parent	0.003	0.100	−0.061	−0.063
		(0.096)	(0.124)	(0.119)	(0.100)
	Grandparents_ no parent	−0.051	0.023	−0.817***	−0.398**
		(0.187)	(0.194)	(0.220)	(0.163)
All - 2SRI	Grandparents_at least one parent	−2.635***	−0.638	−4.350***	−2.837***
		(0.548)	(0.810)	(1.207)	(0.694)
	Grandparents_ no parent	−3.708***	−1.820***	−3.893***	−2.835***
		(0.840)	(0.606)	(0.681)	(0.579)
Boys	Grandparents_at least one parent	−2.910***	−0.895	−4.272**	−1.461
		(0.784)	(1.266)	(1.659)	(0.921)
	Grandparents_ no parent	−2.768**	−1.049	−3.420***	−1.901***
		(1.112)	(0.840)	(0.737)	(0.704)
Girls	Grandparents_at least one parent	−2.498***	−1.970*	−3.473*	−3.282***
		(0.763)	(1.189)	(1.833)	(0.943)
	Grandparents_no parent	−3.300***	−2.988**	−4.771***	−1.992*
		(1.015)	(1.179)	(1.260)	(1.017)
Urban	Grandparents_at least one parent	−1.870**	0.223	−4.110***	−4.177***
		(0.781)	(1.290)	(1.489)	(1.035)
	Grandparents_no parent	−3.698	−0.554	−3.844***	−2.668***
		(7.020)	(1.308)	(1.267)	(0.887)
Rural	Grandparents_at least one parent	−3.161***	−1.076	−0.694	1.533
		(0.731)	(1.520)	(2.544)	(0.937)
	Grandparents_no parent	−3.709***	−1.828**	−2.981***	−2.301***
		(1.000)	(0.812)	(0.997)	(0.856)
Joint significance of IVs (F statistics for full sample)†					
First stage_childcare choice		38.45	45.22	59.29	55.43
First stage_mothers education		26.14	32.85	53.02	44.79
First stage_household income		12.83	12.27	17.30	11.91
Cov(res_n, IV1)‡		−0.023	−0.011	0.013	0.006
Cov(res_n, IV2)		0.000	0.002	−0.001	0.000
Cov(res_n, IV3)		0.114	0.005	−0.006	−0.002
Cov(res_n, IV4)		0.009	0.015	−0.002	0.004
Cov(res_n, IV5)		−0.003	0.000	−0.003	−0.001

Notes: Standard errors with 500 bootstrap replications are reported in parentheses. Original endogenous variables and first-stage residuals are used in two-stage residual inclusion (2SRI) method. Control variables are included.
† F-statistics for the joint significance test of the IVs, P-values are not reported and all <0.01.
‡IV1= 20 percentile wage of female employees in the county; IV2 = dummy of whether the mother is born after 1978; IV3= maternal grandparents' education, IV4=household assets index, IV5=the average proportion of children (under 6 years) left-behind by mothers in the county.
*$P < 0.1$, **$P < 0.05$, ***$P < 0.01$.

$$\log \lambda_i(t|\mathbf{Y}) = \rho_0(t) + \rho_1 G1 + \rho_2 G2 + \rho_3 I + \rho_4 E + \rho_5 X + \rho_6 \hat{\varepsilon}_{G1} + \rho_7 \hat{\varepsilon}_{G2}$$
$$+ \rho_8 \hat{\varepsilon}_E + \rho_9 \hat{\varepsilon}_I + v_{2SRI} \tag{14}$$

where $G1$ is a dummy equal to 1 if the child is cared by a grandparent with at least one parent living at home and 0 otherwise; and $G2$ is a dummy equal to 1 if the child is cared by a grandparent with no parent living at home and 0 otherwise; $\hat{\varepsilon}_{G1}$ and $\hat{\varepsilon}_{G2}$ are,

respectively, the estimated residuals from the multinomial logit model computed as the difference between the observed values and corresponding predicted probabilities, that is $\hat{\varepsilon}_{G1} = G1 - \hat{G1}$; and $\hat{\varepsilon}_{G2} = G2 - \hat{G2}$.[21]

Table 4 reports the results. Overall, grandparental care still has a negative effect on child development. Living with the parent(s) neutralizes this negative effect on children's talking yet being left-behind by both parents would further delay children's development in walking and talking. In subgroup analysis, grandparental care is more harmful to left-behind girls in terms of the onset of walking, talking and counting. Within the rural sample, early childhood development of left-behind children is significantly delayed, while for those who are living with parent(s) the negative effect of grandparental care is mitigated for talking, counting and toilet training. These findings are consistent with previous studies on left-behind children in rural areas in China (Chen *et al.*, 2009; Mu and de Brauw, 2015; Ye and Pan, 2011; Yue *et al.*, 2016).

VII. Conclusions

In this paper, we have used the information on children aged 7 to 59 months from the China Family Panel Studies to examine the impact of different childcare choice on early childhood development, focusing on the timing of four milestone achievements – walking, talking, counting and toilet training. Given the large role of grandparents in looking after young children in contemporary China, our analysis focuses on whether it matters who is the main provider of the daytime childcare – the parents (mainly mothers) or the grandparents. We model the type of childcare provider as an endogenous input into early childhood development using an instrumental variables approach. We also allow for the endogeneity of parental education and household income with our empirical tests rejecting the assumption of exogeneity for these variables and for childcare.

Our main results indicate that compared with parental care, grandparental care delays the achievement of all the four abilities, everything else being equal. In line with these results, parental education overall enhances the speed of acquisition of walking, talking and toilet training but not counting. Household income seems to affect the early development of counting skill and this effect is very strong. Thus, in China, being cared for by parents and having educated parents appears broadly favourable for early childhood development.

Using the subgroup samples, we find some differences between rural and urban areas. The early development of children living in rural areas appears unambiguously impeded by grandparental care, particularly for those who are left behind by both parents. Hetero-geneous effect across gender is not obvious for the whole sample, but left-behind girls are more affected for walking, talking and counting.

On balance, our results imply that having parents as the main carers is beneficial for child development in the first few years, particularly for children in rural areas of China. This echoes the conclusion of previous research on the problem of China's left-behind children being looked after by grandparents, a conclusion that has proved controversial

---

[21] The new IV is also used in the first stage estimations on mothers' education and household income. We also estimate the second stage using the 2SPS method and the results are similar. Joint significance test of the IVs in the first stages and the covariances between each IV and naïve estimates residuals are reported in Table 4.

in the country (Yue *et al.*, 2016, 2017; Zhao, 2017; Zhou *et al.*, 2018). One limitation of our study is that we have only looked at the consequences in terms of the timing of key milestone achievements in early childhood. Whether delaying these achievements has long-term effects on the cognitive achievement of adults is a question for further research.

If we assume that such delays in early childhood development achievements are undesirable, the policy implications of our findings are not straightforward. Part of the controversy surrounding findings of negative consequences of non-parental childcare is the potential inference that women should stay at home, or – in the Chinese context – that migrants should return to the countryside. Such inferences would be simplistic and unrealistic. Fallon, Mazar and Swiss (2017) suggest that maternity leave would benefit child health in developing countries. Studies on early interventions for children's development in developing countries suggest that the most cost-effective approach is to combine support and enhancement of the mothers' resources for care in addition to teaching specific care practices (Engle *et al.*, 1999). In the case of China, these supports may include investment and social support in improving mothers' education and knowledge in care practices as well as time availability. Therefore, improved provision of maternity leave (from birth-related to care-related or more flexible shared parental leave) would allow women to continue their careers without adverse impacts on their children's early development. Nonetheless, given the prevalence of care by grandparents, other policies should also be considered. It may be that educational interventions can offset deficiencies in early childhood development caused by parental absence (Alderman, 2007; Gao *et al.*, 2018). These interventions should consider working with grandparents to enhance the quality of care they provide, particularly in rural areas. Investment in early childhood development in rural China may be crucial for the long-term human capital development of future generations, helping to sustain the country's high economic growth and alleviating the risk of rising inequality.

References

Alderman, H. (2007). 'Improving nutrition through community growth promotion: Longitudinal study of the Nutrition and Early Child Development Program in Uganda', *World Development*, Vol. 35, pp. 1376–1389.

Aunio, P., Aubrey, C., Godfrey, R., Pan, Y. and Liu, Y. (2008). 'Children's early numeracy in England, Finland and People's Republic of China', *International Journal of Early Years Education*, Vol. 16, pp. 203–221.

Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K. and Owen, M. T. (2007). 'Are there long-term effects of early child care?' *Child Development*, Vol. 78, pp. 681–701.

Ben-Porath, Y. (1967). 'The Production of human capital and the life cycle of earnings', *Journal of Political Economy*, Vol. 75, pp. 352–365.

Bernal, R. (2008). 'The effect of maternal employment and child care on children's cognitive development', *International Economic Review*, Vol. 49, pp. 1173–1209.

Bernal, R. and Keane, M. P. (2011). 'Child care choices and children's cognitive achievement: the case of single mothers', *Journal of Labor Economics*, Vol. 29, pp. 459–512.

Biringen, Z., Emde, R. N., Campos, J. J. and Appelbaum, M. I. (1995). 'Affective reorganization in the infant, the mother, and the dyad: the role of upright locomotion and its timing', *Child Development*, Vol. 66, pp. 499–514.

Del Boca, D., Piazzalunga, D. and Pronzato, C. (2018). 'The role of grandparenting in early childcare and child outcomes', *Review of Economics of the Household*, Vol. 16, pp. 477–512.

Bowlby, J. (1951). 'Maternal care and mental health', *Bulletin of the World Health Organization*, Vol. 3, pp. 355–534.

Chen, X., Huang, Q., Rozelle, S., Shi, Y. and Zhang, L. (2009). 'Effect of migration on children's educational performance in rural China', *Comparative Economic Studies*, Vol. 51, pp. 323–343.

Cheng, K. M. (1986). 'China's recent education reform: the beginning of an overhaul', *Comparative Education*, Vol. 22, pp. 255–269.

Clearfield, M. W. (2011). 'Learning to walk changes infants' social interactions', *Infant Behavior and Development*, Vol. 34, pp. 15–25.

Cleary, S. (1999). 'The relationship between firm investment and financial status', *The Journal of Finance*, Vol. 54, pp. 673–692.

Cox, D. R. (1972). 'Regression models and life-tables', *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 34, pp. 187–220.

Cunha, F. and Heckman, J. (2007). 'The technology of skill formation', *American Economic Review*, Vol. 97, pp. 31–47.

Dong, X. and Zhao, Y. (2017). *Care Economy, Gender and Inclusive Growth in Post-Reform China?: How Does Unpaid Care Work Affect Women's Economic Opportunities and Gender Equality?* GrOW Working Paper Series GWP-2017-06.

Du, F. and Dong, X. (2013). 'Women's employment and child care choices in urban China during the economic transition', *Economic Development and Cultural Change*, Vol. 62, pp. 131–155.

Engle, P. L., Menon, P. and Haddad, L. (1999). 'Care and nutrition: concepts and measurement', *World Development*, Vol. 27, pp. 1309–1337.

Fallon, K. M., Mazar, A. and Swiss, L. (2017). 'The development benefits of maternity leave', *World Development*, Vol. 96, pp. 102–118.

Felfe, C., Nollenberger, N. and Rodríguez-Planas, N. (2015). 'Can't buy mommy's love? Universal child-care and children's long-term cognitive development', *Journal of Population Economics*, Vol. 28, pp. 393–422.

Gao, Q., Wang, H., Mo, D., Shi, Y., Kenny, K. and Rozelle, S. (2018). 'Can reading programs improve reading skills and academic performance in rural China? ' *China Economic Review*, Vol. 52, pp. 111–125.

Ghassabian, A., Sundaram, R., Bell, E., Bello, S. C., Kus, C. and Yeung, E. (2016). 'Gross motor milestones and subsequent development', *Pediatrics*, Vol. 138, p. e20154372.

Goodfellow, J. and Laverty, J. (2003). *Grandcaring: Insights into Grandparnts' Experiences as Regular Child Care Providers*, Early Childhood Association, Waston, ACT.

Grantham-McGregor, S. M., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L. and Strupp, B. (2007). 'Developmental potential in the first 5 years for children in developing countries', *Lancet*, Vol. 369, pp. 60–70.

Grantham-McGregor, S. M., Fernald, L. C. H., Kagawa, R. M. C. and Walker, S. (2014). 'Effects of integrated child development and nutrition interventions on child development and nutritional status', *Annals of the New York Academy of Sciences*, Vol. 1308, pp. 11–32.

Gregg, P., Washbrook, E., Propper, C. and Burgess, S. (2005). 'The effects of a mother's return to work decision on child development in the UK', *The Economic Journal*, Vol. 115, pp. F48–F80.

Hansen, K. and Hawkes, D. (2009). 'Early childcare and child development', *Journal of Social Policy*, Vol. 38, pp. 211–239.

Hausman, J. A. (1978). 'Specification tests in econometrics', *Econometrica*, Vol. 46, pp. 1251–1271.

Irwin, L. G., Siddiqi, A. and Hertzman, C. (2007). *Early Child Development?: A Powerful Equalizer, Final Report for the World Health Organization's Commission on the Social Determinants of Health*, March 2007.

Jia, N., Dong, X. and Song, Y. (2018). 'Paid maternity leave and breastfeeding in urban China', *Feminist Economics*, Vol. 24, pp. 31–53.

Leibowitz, A. (1974). 'Home investments in children', *Journal of Political Economy*, Vol. 82, pp. S111–S131.

Leonard, H. C., Bedford, R., Pickles, A. and Hill, E. L. (2015). 'Predicting the rate of language development from early motor skills in at-risk infants who develop autism spectrum disorder', *Research in Autism Spectrum Disorders*, Vol. 13–14, pp. 15–24.

Li, B., Adab, P. and Cheng, K.K. (2015). 'The role of grandparents in childhood obesity in China - evidence from a mixed methods study', *International Journal of Behavioral Nutrition and Physical Activity*, Vol. 12, pp. 91.

Lu, T. (2013). *China's New Workers: Lost or Arising*, Legal Publishing House, Beijing.

Luby, J., Belden, A., Harms, M., Tillman, R. and Barch, D. M. (2016). 'Preschool is a sensitive period for the influence of maternal support on the trajectory of hippocampal development', *Proceedings of the National Academy of Sciences of United States of America*, Vol. 113, pp. 5742–5747.

Luxem, M. and Christophersen, E. (1994). 'Behavioral toilet training in early childhood: research, practice, and implications', *Journal of Developmental and Behavioral Pediatrics*, Vol. 15, pp. 370–378.

Magnuson, K. A., Kelchen, R., Duncan, G. J., Schindler, H. S., Shager, H. and Yoshikawa, H. (2016). 'Do the effects of early childhood education programs differ by gender? A meta-analysis', *Early Childhood Research Quarterly*, Vol. 36, pp. 521–536.

Mo, X., Xu, L., Luo, H., Wang, X., Zhang, F. and Gai Tobe, R. (2016). 'Do different parenting patterns impact the health and physical growth of 'left-behind' preschool-aged children? A cross-sectional study in rural China', *European Journal of Public Health*, Vol. 26, pp. 18–23.

Mu, R. and de Brauw, A. (2015). 'Migration and young child nutrition: evidence from rural China', *Journal of Population Economics*, Vol. 28, pp. 631–657.

National Research Council (2009).*Mathematics Learning in Early Childhood: Paths Toward Excellence and Equity*, The National Academies Press, Washington, DC.

Oudgenoeg-Paz, O., Volmsagean, M. C. J. M. and Leseman, P. P. M. (2012). 'Attainment of sitting and walking predicts development of productive vocabulary between ages 16 and 28 months', *Infant Behavior and Development*, Vol. 35, pp. 733–736.

Rescorla, L. (1989). 'The language development survey', *Journal of Speech and Hearing Disorders*, Vol. 54, pp. 587–599.

Stumbitz, B., Lewis, S., Kyei, A. A. and Lyon, F. (2018). 'Maternity protection in formal and informal economy workplaces: the case of Ghana', *World Development*, Vol. 110, pp. 373–384.

Tan, J.-P., Buchanan, A., Flouri, E., Attar-Schwartz, S. and Griggs, J. (2010). 'Filling the parenting gap? Grandparent involvement with UK adolescents', *Journal of Family Issues*, Vol. 31, pp. 992–1015.

Tchetgen Tchetgen, E. J., Walter, S., Vansteelandt, S., Martinussen, T. and Glymour, M. (2015). 'Instrumental variable estimation in a survival context', *Epidemiology*, Vol. 26, pp. 402–410.

Terza, J. V., Basu, A. and Rathouz, P. J. (2008). 'Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling', *Journal of Health Economics*, Vol. 27, pp. 531–543.

The NICHD Early Child Care Research Network (1998). 'Early child care and self-Control, compliance, and problem behavior at twenty-four and thirty-six months', *Child Development*, Vol. 69, pp. 1145–1170.

The World Bank (2018). *World Development Indicators*, Available at: https://databank.worldbank.org/source/world-development-indicators.

UNICEF (1990).*Strategy for Improved Nutrition of Children and Women in Developing Countries*, A UNICEF Policy Review, UNICEF, New York.

Walle, E. A. and Campos, J. J. (2014). 'Infant language development is related to the acquisition of walking', *Developmental Psychology*, Vol. 50, pp. 336–348.

WHO Multicentre Growth Reference Study Group (2006). 'WHO motor development study: windows of achievement for six gross motor development milestones', *Acta Paediatrica*, 95, pp. 86–95.

Wooldridge, J. M. (2014). *Introduction to Econometrics*, Europe, Mi. Andover, Hampshire: Cengage Learning.

Xie, Y. and Hu, J. (2014). 'An introduction to the China family panel studies (CFPS)', *Chinese Sociological Review*, Vol. 47, pp. 3–29.

Ye, J. and Pan, L. (2011). 'Differentiated childhoods: impacts of rural labor migration on left-behind children in China', *The Journal of Peasant Studies*, Vol. 38, pp. 355–377.

Yue, A., Shi, Y., Luo, R., Chen, J., Garth, J., Zhang, J., Medina, A., Kotb, S. and Rozelle, S. (2017). 'China's invisible crisis: cognitive delays among rural toddlers and the absence of modern parenting', *The China Journal*, Vol. 78, pp. 50–80.

Yue, A., Sylvia, S., Bai, Y., Shi, Y., Luo, R. and Rozelle, S. (2016). *The Effect of Maternal Migration on Early Childhood Development in Rural China*, Available at: https://ssrn.com/abstract=2890108.

Zhao, H. (2017). '*Cognitively Delayed Rural Children Are an Invisible Crisis*', China Daily, Available at: http://www.chinadaily.com.cn/opinion/2017-09/28/content_32581850.htm.

Zhou, M., Sun, X., Huang, L., Zhang, G., Kenny, K., Xue, H., Auden, E. and Rozelle, S. (2018). 'Parental migration and left-behind children's depressive symptoms: estimation based on a nationally-representative panel dataset', *International Journal of Environmental Research and Public Health*, Vol. 15, pp. 1065.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix A: Variables, First Stage Results and Robustness Checks