DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 35, ARTICLE 47, PAGES 1373–1410
PUBLISHED 25 NOVEMBER 2016
http://www.demographic-research.org/Volumes/Vol35/47/
DOI: 10.4054/DemRes.2016.35.47

Research Article

‘Motherhood penalty’ and ‘fatherhood premium’? Fertility effects on parents in China

Zheng Mu

Yu Xie

This publication is part of the Special Collection on “Domestic Division of Labour and Fertility Choice in East Asia,” organized by Guest Editors Ekaterina Hertog and Man-Yee Kan.

Contents

Demographic Research: Volume 35, Article 47
Research Article

'Motherhood penalty' and 'fatherhood premium'? Fertility effects on parents in China

Zheng Mu[1]
Yu Xie[2]

Abstract

BACKGROUND
Many previous empirical findings on 'motherhood penalty' and 'fatherhood premium' remain inconclusive due to potential selection biases. China's regional variation in exemptions to the one-child policy enables us to use the gender of the first child as a powerful instrumental variable (IV) in identifying the gendered fertility effects.

OBJECTIVE
We aim to estimate the causal effects of fertility on fathers' and mothers' various outcomes in China.

METHODS
Using the IV approach, this paper examines the gender-specific fertility effects on parents' time use, income, and subjective well-being, using data for 2010 from the China Family Panel Studies.

RESULTS
Results show that while fathers spend more time at work and less time taking care of family members with more children, mothers report better subjective well-being. Moreover, fathers gain self-confidence in both their careers and the future, and mothers are happier, more satisfied with life and report better social ability.

CONCLUSIONS
Our findings do not directly support the gendered fertility effects on parents. However, the differential fertility effects on specific domains for mothers versus fathers are consistent with household specialisation. By interpreting this conclusion within the context of China's one-child family planning policy, our research suggests that parents would do better if the one-child policy were abolished – i.e., if parents were allowed to have more children.

[1] Corresponding author. Department of Sociology and Centre for Family and Population Research, National University of Singapore, Singapore. E-Mail: socmuz@nus.edu.sg.
[2] Department of Sociology and Princeton Institute for International and Regional Studies, Princeton University, Princeton, New Jersey, USA.

CONTRIBUTION
The unique policy setting in China affords us the methodological opportunity to study the true causal effects of fertility on parents, which has crucial implications for China’s new two-child policy era since October 2015.

1. Introduction

Social scientists have long been interested in the relationship between fertility and employment outcomes (Angrist and Evans 1998; Goldin 1995; Gough and Noonan 2013). In theory, given the constraint on a person’s total time, it seems that childbearing should have negative consequences for labour force outcomes (e.g., Budig and England 2001; Glauber 2007; Goldin 1995; Gronau 1988; Korenman and Neumark 1992). However, it has been hypothesised that the effects of fertility on employment are highly gendered among married couples – negative for mothers but positive for fathers – in that, once they have children, women tend to shift more of their time from paid work to childrearing activities, while men tend to direct greater effort toward bread-winning activities (Becker 1981, 1985; Glauber 2007, 2008; Killewald and Gough 2013). The hypothesised negative effect of fertility on mothers’ labour force outcomes is called ‘motherhood penalty’ (Angrist and Evans 1998; Glauber 2008; Harkness and Waldfogel 2003; Hochschild and Machung 1989; Joshi and Newell 1989; Lundberg and Rose 2000; Neumark and Korenman 1994; Noonan 2001; Waldfogel 1997, 1998a, 1998b), while the hypothesised positive effect for fathers is called the ‘fatherhood premium’ (Killewald 2013; Loh 1996; Lundberg and Rose 2000). Past research using survey data, mostly in the United States, has yielded empirical evidence consistent with these two hypotheses.

Assigning causality to the observed gendered associations between fertility and labour force outcomes from survey data remains controversial, however. Individuals may recognise and account for fertility effects when making childbearing decisions, making fertility endogenous rather than exogenous (Angrist and Evans 1998; Goldin 1995; Gough and Noonan 2013; Schultz 1981). In other words, those who choose to have children may differ from those who do not in observed and unobserved characteristics, such as career motivation, family values, and sense of responsibility – characteristics that relate to both childbearing and labour market outcomes (Budig and England 2001; Gough and Noonan 2013).

Moreover, parenthood is a major part of the family as a social institution and as such can profoundly impact parents in more than purely economic domains. Although the effects of fertility on parents’ labour force outcomes are very important for both the social science and public policy literature, they should not be the sole focus in understanding the consequences of fertility for parents. Specifically, fertility may change parents' lifestyles and perspectives so that having children is associated with improved subjective well-being, that is, the 'subjective premium' (Aassve, Goisis, and Sironi 2011; Baranowska and Matysiak 2011; Billari and Kohler 2009; Hoffman and Hoffman 1973; Kohler, Behrman, and Skytthe 2005; Waite and Gallagher 2000; Margolis and Myrskylä 2011). However, empirical evaluation of the subjective premium hypothesis is also fraught with methodological difficulties stemming from potential endogeneity. That is, those who derive more subjective rewards from childbearing are also more likely to become parents.

In this paper, we capitalise on regional variation in implementation of the one-child policy in contemporary rural China and use the gender of the first child as an instrumental variable (IV) to identify the causal effects of fertility on parents' time use, labour force, and psychological outcomes. For this task, we analyse the newly available data for 2010 from a nationally representative survey – the China Family Panel Studies (CFPS) – matched regionally by the differential local implementation of the one-child policy.

The choice of China as our study site is motivated by high theoretical interest as well as methodological convenience. In the past three decades, China's economy has grown rapidly (Xie 2011). During this period, women's socioeconomic status has improved tremendously, and the traditional Chinese family, with patriarchy at its core, has been significantly eroded (Xie 2013). In particular, gender-oriented specialisation within households has been weakened (Bian, Logan, and Shu 2000; Whyte and Parish 1984; Wolf 1984; Yu and Xie 2011; Zuo and Bian 2001). Fertility has also been very low due to the Chinese government's one-child policy. It is possible that the effects of fertility on parents do not differ between fathers and mothers. However, to the best of our knowledge, there is no systematic research on the gendered fertility effects on China, and we aim for that in this paper. Moreover, China's differential implementation of the one-child policy has also enabled the use of an IV approach. So far, whether and how fertility affects Chinese parents' labour force participation and subjective outcomes remain unknown and merit empirical examinations.

Our analyses address two related research questions: (1) does having more than one child influence the parents' time use, income, and subjective well-being in China? (2) If yes, how are the effects different for fathers and mothers?

2. Theoretical issues and research setting

2.1 'Motherhood penalty' and 'fatherhood premium'

The model of within-household specialisation posits that couples pursue a joint strategy in which they divide labour to maximise the well-being of the household (Becker 1981, 1985). The division of labour, typically with the husband specialising in the labour market and the wife specialising in home production, is based on the comparative advantages of the spouses in each realm. Traditional socialisation is highly gendered, encouraging men to develop skills for the labour market and women to become capable housewives (Becker 1981, 1985), and the labour market seems to support this specialisation as well, given that employed women have historically earned less than employed men (Bianchi 1994; Blau 2012; Corcoran and Courant 1987; Oppenheimer 1997; Smock, Manning, and Gupta 1999). Household specialisation serves as the main causal explanation for women's 'motherhood penalty' and men's 'fatherhood premium' concerning the effects of childbearing on labour market outcomes (Budig and England 2001; Glauber 2008; Gough and Noonan 2013; Killewald and Gough 2013; Noonan and Corcoran 2004; Waldfogel 1997). Its causal impact may emanate from three possible mechanisms.

First, gendered investment in, and accumulation of, human capital in the labour market predicts differential effects of fertility for mothers and fathers. When women specialise by assuming the primary childrearing role, they spend less time in the labour force, accumulating less employment experience and ultimately being paid less (Becker 1981, 1985; Polachek 1985). Men with children, on the other hand, may be motivated more by their specialised provider role to accumulate greater human capital in the labour market.

Second, household specialisation may affect the kinds and extent of effort men and women are able or willing to put into their work, and thus the type of employment they choose. Women who assume primary childrearing responsibilities may have less energy and time for labour market activities than women without children (Becker 1981, 1985). Fathers, on the other hand, may put more effort into their careers in their roles as main earners than do their childless counterparts (Becker 1981, 1985). Moreover, some mothers may 'institutionalise' this division of household work by choosing more flexible and accommodating jobs, which usually offer lower compensation. This trade-off between job flexibility and compensation, known as the 'compensating differential,' supplements other explanations of mothers' labour force disadvantage (England 1992; Filer 1985).

Finally, these gendered differences in accumulation of marketable human capital and in a choice of demanding jobs may signal to the labour market that mothers tend to be less productive employees than either fathers or women with fewer/no childrearing responsibilities. Employers may practise statistical discrimination against mothers, paying them less than non-mothers or fathers for the same types of jobs or assigning them to lower-paying jobs (Arrow 1972, 1973; Becker 1957; Phelps 1972). Fathers, however, do not suffer parallel employment discrimination (Arrow 1972, 1973; Becker 1957; Phelps 1972).

While these mechanisms may indeed link fertility to mothers' labour market disadvantage, demonstrating causality remains elusive. Some research indicates that observed fertility effects may result from selection bias (Angrist and Evans 1998; Budig and England 2001; Gough and Noonan 2013; Jacobsen, Pearce, and Rosenbloom 1999; Korenman and Neumark 1992; Lundberg and Rose 2000; Miller 2011; Waldfogel 1997). Individuals who decide to become parents may differ from nonparents in characteristics that relate to labour market outcomes, such as career aspirations, work commitment, family values and sense of responsibility (Budig and England 2001; Gough and Noonan 2013). Also, individuals may make decisions about their fertility behaviours based on their labour market and financial situations (Angrist and Evans 1998; Gough and Noonan 2013). For example, when the labour market condition for a woman of childbearing age becomes unfavourable, she may be more inclined to take up the role of homemaker and mother. Conversely, men may become more motivated to have children if they achieve employment and financial success. In short, a potential endogeneity threat suggests that causality may operate in the other direction: from labour market outcomes to fertility decisions.

Most studies addressing selection bias have either directly controlled for possible differences between parents and non-parents or have exploited a longitudinal dataset structure with fixed-effects models, which eliminates between-individual variation that stays stable over time (Becker 1985; Blank 1990; Budig and England 2001; Gough and Noonan 2013; Hill 1979; Korenman and Neumark 1992; Lundberg and Rose 2000; Waldfogel 1997). However, with the former method, identifying all relevant observed differences between parents and nonparents, or between parents with more and fewer children, is a difficult empirical task. Studies using this method are still subject to the criticism that additional relevant factors remain unobserved and thus uncontrolled. With the latter method, researchers need to assume that potential confounders threatening causal inference are fixed over time. Thus, an alternative method for dealing with potential selection bias in establishing causality is the instrumental variable (IV) approach (Angrist and Evans 1998), which requires the use of a predicting variable that is exogenous to the outcome variable. This condition can usually be satisfied only in natural experiments or through rigorous treatment designs (e.g., Angrist and Evans 1998; Jacobsen, Pearce, and Rosenbloom 1999; Miller 2011).

Finally, most studies of fertility effects have focused on employment and financial outcomes (Angrist and Evans 1998; Budig and England 2001; Glauber 2007, 2008; Hill 1979; Killewald 2013; Killewald and Gough 2013; Waldfogel 1997). However, as childbearing and childrearing are such important life events, they influence well-being in a broader sense, especially regarding subjective outcomes (Waite and Gallagher 2000). Recently, more and more researchers in family studies have been paying attention to the fertility effects on subjective outcomes, and most of them find subjective premiums (Aassve, Goisis, and Sironi 2012; Baranowska and Matysiak 2011; Billari and Kohler 2009; Hoffman and Hoffman 1973; Hoffman, Thornton, and Manis 1978; Kohler, Behrman, and Skytthe 2005; Kravdal 2013; Margolis and Myrskylä 2011; Waite and Gallagher 2000). In Hoffman, Thornton, and Manis's (1978) work, the authors divide the subjective fertility premiums into nine major categories: primary group ties and affection, stimulation and fun, expansion of the self, adult status and social identity, achievement and creativity, morality, economic utility, power and influence, and social comparison. This categorisation reveals the multitude of potential mechanisms that link the parents' fertility behaviours to their subjective well-being.

2.2 The Chinese context

In recent decades Chinese society has undergone dramatic social change (Hauser and Xie 2005; Xie and Hannum 1996; Xie et al. 2013). Two of the most significant changes are women's improved social status (Hannum 2005; Treiman 2013; Wu and Song 2010; Wu and Zhang 2010; Zhang, Hannum, and Wang 2008) and the evolution of China's one-child policy (Greenhalgh 2008; Gu et al. 2007; Guo, Liu, and Song 2002). These two changes make China an ideal research setting in which to examine fertility effects, from both theoretical and methodological perspectives.

Theoretically, the Communist Revolution and the government's enthusiastic promotion of gender equality should have significantly improved women's social status (Lavely et al. 1990). During the decades following the 1949 Revolution, Communist ideology regarding gender equality was zealously promoted, highlighting women's parity with men (Meisner 1999; Parish 1981; Whyte 2010; Yu and Xie 2013) and popularising the slogan 'women hold up half the sky' (Maurer-Fazio, Rawski, and Zhang 1999). In the spheres of politics and work life, the Chinese constitution guarantees women equal rights with those of men in all respects and specifically endorses the policy of 'same work, same pay' (Maurer-Fazio et al. 1999; Zuo and Bian 2001). As proposed by Goldscheider, Bernhardt, and Lappegård (2015), women's more active involvement in labour force participation is only the first step of the 'gender revolution' which aims to improve women's social standing, while the second step –

men joining women in the private sphere of the family – is not yet pervasive. However, in China, even this second step has been widely promoted. Specifically, in the sphere of family life, in 1950, China instituted the Marriage Law, which formally legalised freedom of choice in marriage and explicitly protected wives' rights and interests, making them equal to those of their husbands (China Administration Council 1950: Item 5; Zuo and Bian 2001). These ideological and policy changes have significantly enhanced women's social standing and economic status in contemporary China (Hannum 2005; Lavely et al. 1990; Song 2009; Zhang, Hannum, Wang 2008). Women's educational attainment has gradually caught up with that of men (Treiman 2013; Wu and Song 2010: Table 2; Wu and Zhang 2010), the gender gap in income and labour force participation has narrowed, and women have started to assume premium positions that had previously been dominated by men (Meng 1993; Parish and Busse 1998). Home life has not been immune to these shifts, with household gender inequality and within-household specialisation gradually declining since the Revolution (Bian, Logan, and Shu 2000; Whyte and Parish 1984; Wolf 1984; Yu and Xie 2011; Zuo and Bian 2001). These circumstances may have changed the mechanisms by which the 'motherhood penalty' and the 'fatherhood premium' are thought to operate.

Regarding the outcomes being studied, China's policy background may make the link between fertility behaviours and subjective well-being especially pertinent. China began its nation-wide family planning programme as early as 1973. In the beginning, the policy was simply a general promotion of 'later, sparser, and fewer' ('*wan, xi, shao*') fertility behaviour. In 1980, however, the policy was formalised into a restriction allowing all couples to have only one child. Later, as an adjustment to China's pro-natal culture and its traditional preference for male heirs, the government amended its family planning policy to allow some couples to have a second child under certain conditions, the major one being that the couple's first child was a girl (Peng 1997). This version of the policy has been applied since 1988 (Guo, Liu, and Song 2002; Peng 1997). Since China's family planning policy was implemented mandatorily, the realised fertility level may not reflect individuals' real preferences regarding family size. This constraint on individuals' capability to realise their preferences is believed to negatively influence their subjective well-being (Eibach and Mock 2011; Margolis and Myrskylä 2011; Nelson et al. 2013; Wang, Jing, and Zhang 2013; White and Dolan 2009). Therefore, compared to those who can have only one child, those individuals who are able to have more children may have realised their preferences to a larger extent and thus may have better subjective well-being.

Methodologically, regional variations in the implementation of China's one-child policy based on the gender of the first child affords us the opportunity to study fertility effects under these new circumstances regarding the gender power structure in the restructured modern household. Essentially, we have an ideal natural experiment in which to implement the IV strategy in order to identify the causal effects of fertility on mothers and fathers. As previously mentioned, the one-child policy for married couples in China was formalised in 1980 (Greenhalgh 2008; Guo, Liu, and Song 2002). However, this initial version of the policy was eventually deemed too drastic and inflexible, ignoring the potential heterogeneities in fertility intentions and behaviours across regions and, in particular, across the urban-rural divide (Guo, Liu, and Song 2002). For example, Chinese society has historically maintained a patriarchal and patrilineal family system which values larger family size and favours sons over daughters (Thornton and Lin 1994; Xu, Ji, and Tung 2000; Whyte 2003). These traditional family values have been more strongly maintained in rural areas than in urban ones (Guo, Liu, and Song 2002). Accordingly, in 1988, the one-child policy was tailored to accommodate these contextual specificities (Guo, Liu, and Song 2002). The most significant adjustment was that in specified – primarily rural – areas the policy was made flexible according to the gender of the first child: if the first child was a girl, the parents were allowed to have a second (Gu et al. 2007; Guo, Liu, and Song 2002).[3] Therefore, in these areas, parents whose first child is a girl are substantially more likely to have one or more additional children.[4] Since gender at birth is virtually random, especially when having a girl as the first child does not prevent parents from having a second child,[5] whether the first child is a girl or a boy is an excellent IV for additional childbearing among parents with at least one child, and allows us to evaluate the causal effect of fertility free from potential selection biases in traditional regression analyses with observational data.

3. Data and methods

This study uses the instrumental variable (IV) approach to examine the gender-specific effects of fertility on parents' time use, income, and subjective well-being. Our data source is a 2010 sample from the nationally representative China Family Panel Studies (CFPS). The CFPS covers a wide range of information on individuals' social and economic activities, family backgrounds, and subjective outcomes. Specifically, we use the adult sample for the parents' information and derive the children's, the spouses', and the grandparents' information by linking family members within a sampled household.

[3] For details of the policy, please refer to Appendix Table A-1.
[4] This is supported by the results in Table 1.
[5] For discussion on the validity of this assumption, see Section 3.4.

3.1 Analytical sample

First, based on the adult sample and the linked information from family members, we restrict the adult sample to those with children. Then, to secure the basic validity of the analysis, we keep only those who have eligible values for all the independent variables, including the endogenous variables, the instrumental variable, and the control variables. To maximise the relevance and comparability of the outcome variables across individuals, we further restrict the sample to those aged 20 to 50, the prime working ages, and to those who have not yet retired.

Then, regarding childbearing behaviour, we restrict the sample to those parents whose first child is 18 or younger – i.e., all of whose children still require intensive parental care. Restricting our analyses to the parents of children in this age range also ensures that the respondents' childbearing outcomes occurred during the time of the exemption policy, which started in 1988. To maximally ensure that the focal couples are the biological parents of the children, we further restrict our study to couples in their first marriages.

Finally, to make the exemption policy relevant, we include respondents who live in provinces where a second child is allowed if the first child is a girl, have rural residential registration status, and are ethnic majority Han. The ethnic restriction is necessary because minorities living in rural areas are generally allowed to have at least two children (China State Ethnic Affairs Commission 1999). To handle missing values, we carried out multiple imputations with the independent variables as listed in Section 3.3, Variables. These restriction and imputation procedures leave us with an analytic sample of 1,123 fathers and 868 mothers.

In Appendix Table A-1 we show that several other conditions besides the gender of the first child also may trigger the exemption policy. However, we apply no additional sample restrictions given our lack of relevant information on these other potential conditions. As a robustness check, we experimented with different versions of sample restrictions using all the relevant information from the CFPS dataset, and the results (not shown) remained highly consistent.

3.2 Instrumental Variable (IV) approach

The instrumental variable (IV) approach is among the most powerful methods for dealing with the selection bias issue in establishing causality. An IV affects the endogenous explanatory variable while not affecting the outcome variable other than through its effect on the key explanatory variable. We can estimate the causal effect by the indirect least squares estimator (ILS). Let us denote the parental outcome by Y, having more than one child by *X*, and having a female first child by *Z*. The reduced-form, linear model gives us the total influence of having a girl first on *Y*:

$$Y_i = \Pi_0 + \Pi_1 Z_i + v_i \text{, where } \Pi_1 = \frac{\partial Y}{\partial Z} \tag{1}$$

We are interested in this reduced form model insofar as it gives us statistical leverage to estimate a different parameter of interest – the coefficient indicating the fertility effect on *Y* in the following structural equation:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \text{, where } \beta_1 = \frac{\partial Y}{\partial X} \tag{2}$$

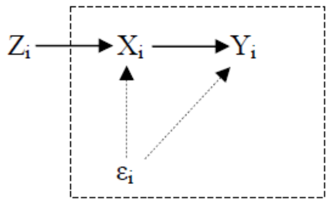
Combining equations (1) and (2) gives the following relationship:

$$\Pi_1 = (\frac{\partial X}{\partial Z}) * (\frac{\partial Y}{\partial X}) = (\frac{\partial X}{\partial Z}) * \beta_1 \tag{3}$$

When the fertility decision is endogenous – that is, when *X* is endogenous to *Y* – we cannot directly estimate β_1 in (2). For example, it is possible that family-oriented parents may tend both to have more children and to earn more than do less family-oriented parents so that selection bias threatens the estimations of causal effects. In this analysis we instead estimate the fertility effect indirectly using an IV.

Figure 1 shows the mechanisms for the IV approach versus the structural model (2) (in box). The structural model depicts the potential for some unobserved characteristics in the residual ε to have an impact on both X and Y. As above, ε could be one's family orientation, which may affect both the likelihood of having multiple children (X) and the amount of the potential outcome (Y), resulting in a selection bias in the observed relationship between X and Y. In the meantime, given its nearly random occurrence, the instrumental variable 'gender of the first child' (Z) is exogenous to Y but highly predictive of the probability of having more than one child. If having more than one child (X) has a causal impact on Y, Z also covaries with Y. Therefore, if we observe a significant association between gender of the first child (Z) and Y, we can indirectly estimate the causal effect of having more than one child (X) on Y under the two-fold assumption that Z serves as a valid IV: (1) Z affects X; (2) Z affects Y only indirectly through X. Assumption (2) is called the 'exclusion restriction.'

Figure 1: Illustration of the IV mechanisms



Specifically, we can obtain the first component in (3) by estimating the following model (also called the first-stage model in a two-stage least squares estimation):

$$X_i = \theta_0 + \theta_1 Z_i + u_i \text{, where } \theta_1 = \frac{\partial X}{\partial Z} \tag{4}$$

The IV estimate is then given by the ratio of the reduced-form estimate in (1) to the coefficient from (4):

$$\beta_1 = \frac{\partial Y}{\partial X} = \frac{\Pi_1}{\theta_1} \tag{5}$$

Based on this estimation procedure, and assuming that gender of the first child is randomly assigned, we can then purge X of the selection bias and obtain an estimate of the causal effect of fertility on parental outcomes. Allowing for heterogeneous treatment effects, we may also interpret β_1 as a 'local average treatment effect' (LATE), specific to the instrument, Z (Angrist, Imbens, and Rubin 1996). In this case β_1 estimates the average effect of X on Y for individuals whose fertility has been influenced by the gender of their first child. To be more concrete, since Π_1 only captures the amount of treatment effect for those whose fertility has been affected by the gender of their first birth, we need to attribute the overall reduced-form estimate to the proportion affected, θ_1, so as to obtain the LATE for the group being affected by the IV.

The results also present coefficients based on ordinary least squares (OLS) estimation for the purposes of comparison. Based on the comparison between the results from ILS and OLS, we can directly observe the differences between causal inference and regular analysis using observational data.

3.3 Variables

Instrumental variable: gender of the first child. This binary variable, coded 0=male and 1=female, is randomly assigned and highly correlated with the tendency to have more children among those affected by the exemption policy.

Endogenous independent variables: fertility level. We use two measures of fertility level. ‘Having more than one child’ is a binary variable coded 0=having one child and 1=having more than one child. ‘Number of children’ is a continuous variable for the total number of children.

Outcome variables: We capture three domains of outcome variables. We use two time use variables: hours worked per month in 2009 and hours taking care of family members in the prior month. To make the measure of labour force participation more reliable, working hours are calculated as hours worked per day multiplied by days worked per month in 2009. Given its variability across individuals, we use its logged form in our analyses. For the second time-use variable, we total for the prior month the average daily hours taking care of family members both during weekdays and over the weekend. To make the family care variable comparable in scale to the labour force participation variable, we then multiply the weekly estimate by four and take its natural logarithm in the analysis. The income variable is measured as personal income in the prior month. Since income varies greatly across individuals in the sample, we use its logged form in the analysis. The subjective well-being variables are measured in six areas on a scale from 1 to 5, with a larger number indicating greater well-being. The six self-rated areas are happiness, life satisfaction, self-confidence in career, self-confidence in the future, self-rated quality of social relationships, and self-rated social ability. In addition, we compute a composite scale, the average of the six self-ratings, to indicate overall subjective well-being.

Control variables: In principle, control variables are included in an IV analysis for two purposes. First, they may be necessary for identification if the IV is not purely random. Second, they may help improve estimation precision if they are strong predictors of outcomes net of the endogenous independent variable, that is, the main treatment variable. In this paper, as we will discuss in more detail in Section 3.4, Limitations, our IV, gender of the first child, is largely randomly assigned. Therefore, we include the following control variables to improve estimation precision and control for the observed heterogeneity that may influence both the independent and dependent variables.

Specifically, we include a rich set of control variables that capture: work in an agricultural industry (binary; 0=no), migrant status (0=no), education (continuous; in years), age (years), age at first birth (years), age gap between the oldest and youngest child (years), living with the youngest child (0=no), living with spouse (0=no), and living with the child(ren)’s grandfather (0=no, 1=yes and grandfather is below age 70, and 2=yes and grandfather is age 70 or above) or grandmother (0=no, 1=yes and grandmother is below age 70, and 2=yes and grandmother is age 70 or above). Specifically, for age and age at first birth, we want to use these two variables to indicate the parent's work experience and their positions in career pathways, currently and at the time of first birth. Moreover, the difference between the two will be the age of the oldest child. Since we do not expect interpretations from coefficients on the control variables, and the inclusion of age and age at first birth has already captured the variation attributable to the age of the first child, we do not directly include the child's age. We then further include age gap between the oldest and youngest child to capture variations due to the age of the youngest child as well as the likelihood of the oldest child taking care of the youngest child. Also note that, for grandparents' coresidence, we make separate categories based on the grandparent's age so as to specify the direction of the care. For example, living with grandparents in their 50s or 60s is often for the purpose of 'grandparenting,' while living with grandparents in their 70s or 80s is mostly for taking care of the elderly. We are aware that age gap between the oldest and youngest child and variables indicating coresidence status with the youngest child and grandparents may be affected by the fertility level, which may lead to biased estimates of the fertility effects. However, those variables are also strong predictors of several outcomes, including work hours, hours spent taking care of family members, income, and subjective well-being. This is particularly true when we consider that those variables may indicate the couple's other caregiving responsibilities, aside from caring for the youngest child, and their attitudinal preference for larger families. Therefore, excluding those variables from the models may lead to even more biased results due to omissions of the theoretical nuances captured.

3.4 Limitations

We are aware of the limitations of the methodology. First and foremost, the gender of the first child as a valid IV may be challenged. For example, when the parents' first child is a boy, they may strategically work harder so as to earn more in anticipation of the financial burdens involved in preparing for the boy's future wedding and household establishment. In China, especially in rural China, these costs are traditionally borne by the groom's parents. That is, the gender of the first child may affect the outcomes directly rather than only through affecting fertility, thus violating the exclusion restriction. However, since our analytical sample includes parents of relatively young children, this may not yet be very relevant. It may also be asked whether or not the gender of the first child is randomly assigned, given the increasing prevalence of sex-selective abortions (Chu 2001). However, research suggests that in China this approach to sex-selection is significantly more prevalent for second or higher-parity births than for first births (Gu et al. 2007). This may be especially true for our analytic sample, in which most parents were allowed to have a second child if their first child was a girl, reducing their motivation to use sex-selective abortion for the first pregnancy. The sample distribution itself also suggests randomness in first child gender since, for fathers and mothers respectively, 49.64% and 49.65% of their first children were girls. Furthermore, we can only capture the treatment effects of higher-parity births using the gender of the first child as the IV. That is, how parents are affected by their transitions from one child to two children. In the meantime, effects of the transitions from no child to one child cannot be estimated. The nature of the IV method imposes this limitation. In the future, our findings may be complemented by further studies estimating the general fertility effects, using alternative methods such as fixed-effects models.

Second, since the analyses are conducted on a rural sample where farming is typically family-based, the income outcome variables may be a poor measure for labour force outcomes. This might be the reason why we do not detect a positive effect of fertility on fathers' income even while fertility lengthens fathers' working hours in the following analysis. Moreover, this may affect the generalisability of the estimated treatment effects, as those who are eligible for, and who have actually utilised, the policy exemptions are mostly rural, tend to be socioeconomically disadvantaged, and may hold a strong preference for more children and male descendants. However, due to the research design of IV, we have to restrict the sample to this specific group in this paper. The recently implemented (October 2015) universal 'two-child policy' in China may in future provide us with a methodological opportunity to estimate the fertility effects on a sample representative of the general population of China.

Finally, due to the long list of restrictions, we are left with a relatively small sample size for selected rural areas for our analyses.

4. Results

4.1 Validity of the instrumental variable

Tables 1a and 1b show differences in means, by gender of the first-born child, for the two endogenous fertility variables and the outcome variables, separately for fathers (Table 1a) and mothers (Table 1b). The differences in means for the fertility variables comprise the θ_1 in Equation (5), indicating the strength of the association between the instrumental variable and the endogenous fertility variables. As shown for the father sample (Table 1a), about half of all fathers whose first child was a boy went on to have a second child; whereas about 70% of fathers whose first child was a girl had a second child. Correspondingly, the number of children in boy-first families tends to be smaller than that in girl-first families – 1.55 children versus 1.91, respectively, in the father sample. The mother sample (Table 1b) shows similar patterns in fertility by gender of the first-born child, with 51% of boy-first mothers having a second child compared to 73% of girl-first mothers, and boy-first mothers having fewer children on average – 1.56 compared to 1.94 for girl-first mothers. Moreover, among both fathers and mothers and for both the endogenous fertility variables, θ_1 remains significantly positive. These results confirm the validity of gender of the first child as an instrument for the two fertility variables.

The differences in means for the outcome variables in Tables 1a and 1b comprise the Π_1 component in Equation (5), the reduced-form parameter. For fathers, the directions of our estimates are consistent with household specialisation – that is, having a girl as the first child has a positive reduced-form relationship with working hours, personal income, and subjective well-being, and a negative association with hours spent caring for family members. However, only the positive association with working hours is statistically significant, with girl-first fathers working around 13% (exp (0.12)-1) more hours compared to boy-first fathers. For mothers, the estimates are also consistent with the specialisation theory in that girl-first mothers tend to work less, spend more time caring for family, earn less, and have better subjective well-being than boy-first mothers. However, only the subjective well-being association is statistically significant, with girl-first mothers enjoying a 0.09-point well-being premium over boy-first mothers on a scale from 1 to 5. Note that the reduced-form estimates need to be adjusted by the magnitude of IV effects on fertility behaviours – that is, the proportion affected – to derive the LATE for the subpopulation being affected by the IV. To do that, we take the ratio of estimates in the two panels in Tables 1a and 1b, and the end product is called the ‘Wald estimates.’ Since Wald estimates are similar to the indirect least square (ILS) estimates, both in interpretations and estimate magnitudes, we present the former only in Appendix Tables A-3a and A-3b.

Table 1a: Difference in means for fertility (denominator) and outcome variables (numerator), father sample

	Gender of the first child				Difference in (Female minus Male)		
	Male (*N*=566)		Female (*N*=557)				
Variables	Mean	s.d.	Mean	s.d.	Mean		s.e.
Fertility variables							
More than one child (ref.=one child)	0.49	0.50	0.70	0.46	0.21	***	0.03
Number of children	1.55	0.61	1.91	0.79	0.36	***	0.04
Outcome variables							
Time use outcomes							
Logged hours worked per month in 2009	5.18	0.82	5.29	0.54	0.12	**	0.04
Logged hours taking care of family members last month	–0.21	4.01	–0.32	4.00	–0.11		0.24
Income outcome							
Logged personal income last month	4.09	5.43	4.37	5.26	0.29		0.32
Subjective Outcome							
Overall subjective scale	3.78	0.66	3.83	0.66	0.05		0.04

Note: 2010 CFPS. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Overall subjective scale is the average of six subjective scales ranging from 1 to 5 on overall happiness, life satisfaction, self-confidence in career, self-confidence in the future, quality of social relationship, and social ability. Larger numbers indicate more positive ratings. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

Table 1b: Difference in means for fertility (denominator) and outcome variables (numerator), mother sample

	Gender of the first child				Difference in (Female minus Male)		
	Male (*N*=437)		Female (*N*=431)				
Variables	Mean	s.d.	Mean	s.d.	Mean		s.e.
Fertility variables							
More than one child (ref.=one child)	0.51	0.50	0.73	0.45	0.22	***	0.03
Number of children	1.56	0.59	1.94	0.80	0.39	***	0.05
Outcome variables							
Time use outcomes							
Logged hours worked per month in 2009	5.03	0.86	4.99	0.94	–0.04		0.06
Logged hours taking care of family members last month	1.05	3.97	1.21	3.94	0.16		0.27
Income outcome							
Logged personal income last month	1.36	5.78	0.95	5.75	–0.41		0.39
Subjective outcome							
Overall subjective scale	3.77	0.62	3.86	0.65	0.09	*	0.04

Note: 2010 CFPS. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Overall subjective scale is the average of six subjective scales ranging from 1 to 5 on overall happiness, life satisfaction, self-confidence in career, self-confidence in the future, quality of social relationship, and social ability. Larger numbers indicate more positive ratings. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

Table 2: OLS estimates of fertility (denominator) and outcome (numerator) equations

	Gender of the first child (ref.=male)			
Dependent variables	Father (*N*=1,123)		Mother (*N*=868)	
Denominator: fertility variables				
More than one child (ref.=one child)	0.12	***	0.11	***
	(0.02)		(0.02)	
Number of children	0.24	***	0.24	***
	(0.03)		(0.03)	
Numerator: outcome variables				
Time use outcomes				
Logged hours worked per month in 2009	0.11	**	–0.01	
	(0.04)		(0.06)	
Logged hours taking care of family members last month	–0.47	*	–0.40	†
	(0.23)		(0.24)	
Income outcome				
Logged personal income last month	0.28		–0.17	
	(0.29)		(0.30)	
Subjective outcome				
Overall subjective scale	0.07	†	0.11	**
	(0.04)		(0.04)	

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Overall subjective scale is the average of six subjective scales ranging from 1 to 5 on overall happiness, life satisfaction, self-confidence in career, self-confidence in the future, quality of social relationship, and social ability. Larger numbers indicate more positive ratings. All models are estimated with control variables described in Appendix Table A-2. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

Table 2 shows results comparable to those in Table 1, except that they are estimated with control variables. As can be seen for the Stage 1 fertility variables, θ_1 remains significantly positive for all the combinations of fertility variables and parent gender, again supporting the validity of the instrument. For Stage 2 outcome variables, girl-first fathers work 12% (exp(0.11)-1) more hours per month, spend 37% (1-exp(-0.47)) fewer hours caring for family members and enjoy a marginally significant 0.07-point bonus in subjective well-being. Mothers, on the other hand, remain unaffected by fertility in terms of time use and income except for hours caring for family members, but enjoy a 0.11-point subjective well-being premium on a scale from 1 to 5. At first sight, it is surprising that the coefficient on logged hours taking care of family members last month is also negative, although only marginally significant. Then, considering the vague definition of the 'family members' being cared for in this variable, it is not as surprising. It is possible that, after having more children, the mother may spend more time taking care of them and less time taking care of other family members. For fathers, it is possible that they may reduce their care-giving responsibilities to a larger extent. That, in part, may explain the larger significance of the coefficient for fathers.

To further establish the validity of the ILS IV estimates, we conducted a balance check of the control variables across values of the instrumental variable. As can be seen in Appendix Table A-2, for almost all cases the mean of the control variables does not differ significantly by gender of the first child. This check shows that the IV is not only exogenous to the outcome variables, but is also highly exogenous to other independent variables included in the full model, which further demonstrates the unbiased nature of the ILS IV estimates.

4.2 'Fatherhood premium'?

Table 3a presents fertility effects for fathers using ILS estimation and compares these to OLS estimates. As can be predicted by the results presented in Table 2, ILS estimates show that fathers with more children tend to work more hours, spend fewer hours taking care of family members, and report marginally higher subjective well-being than do fathers with fewer children. Specifically, fathers with more than one child work more than twice as many hours as fathers who have only one child, and fathers, in general, increase their working hours by about 62% with each additional child. Fathers with more than one child also spend about 98% less time than one-child fathers caring for family members. Each additional child leads to a reduction in time spent caring for family members by around 86% for fathers. Moreover, fathers with more than one child enjoy a 0.58-point subjective well-being premium; with each additional child, fathers' overall subjective well-being increased by 0.29 points on a 1-to-5 scale. However, coefficients for both time spent with family members and overall subjective well-being are only marginally significant. Moreover, coefficients for logged personal income, whether using having more than one child or the number of children as the endogenous variable, though insignificant, are both positive. The results are consistent with the household specialisation theory, with the exception of no significant result for the logged personal income, which could be due to the vagueness of income measure for rural residents.

Table 3a: OLS and ILS estimates of outcome models, father sample

Dependent variables	Father (*N*=1,123)							
	Time use outcomes				**Income outcome**		**Subjective outcome**	
	Logged hours worked per month in 2009		Logged hours taking care of family members last month		Logged personal income last month		Overall subjective scale	
Estimation methods								
More than one child								
OLS	–0.001		–0.16		–0.39		–0.01	
	(0.06)		(0.36)		(0.44)		(0.06)	
ILS	0.95	*	–3.93	†	2.35		0.58	†
	(0.38)		(2.05)		(2.47)		(0.35)	
Number of children								
OLS	0.02		0.06		–0.17		0.02	
	(0.04)		(0.23)		(0.28)		(0.04)	
ILS	0.48	**	–1.98	†	1.18		0.29	†
	(0.18)		(1.01)		(1.23)		(0.17)	

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Overall subjective scale is the average of six subjective scales ranging from 1 to 5 on overall happiness, life satisfaction, self-confidence in career, self-confidence in the future, quality of social relationship, and social ability. Larger numbers indicate more positive ratings. All models are estimated with control variables described in Appendix Table A-2. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

For almost all the significant ILS results in Table 3a, the OLS counterparts are either in the opposite directions or in the same directions, though with much smaller magnitudes. None of the OLS estimates is significant. For example, for fertility effect on working hours, the ILS estimates are as high as 0.95 and 0.48, while the OLS estimates are respectively –0.001 and 0.02. For fertility effect on income, the ILS estimates are positive, in accordance with the specialisation theory, while the OLS estimates are negative, running counter to it.

Table 3b shows the item-specific fertility effects on the six subjective well-being outcomes for fathers. Fertility was positively related to the two self-confidence variables related to careers and the future. Specifically, fathers with more than one child are more confident in their careers by 1.30 points, with a 0.65-point bump for each additional child, and more confident in the future by 1.09 points, with a 0.55-point rise per additional child. These components of subjective well-being, which reflect fathers' sense of their role in their families' current and future well-being, are supportive of the fatherhood premium in time use, in that the higher self-confidence of fathers with more children may emanate from their more highly developed career orientations. We found no other significant subjective well-being results based on either the ILS or the OLS estimations.

Table 3b: OLS and ILS estimates of subjective outcome models, father sample

	Father (N=1,123)							
	General		**Confidence**				**Social**	
Dependent variables	Self-rated happiness	Life satisfaction	Self-confidence in career		Self-confidence in the future		Self-rated quality of social relationship	Self-rated social ability
Estimation methods								
More than one child								
OLS	–0.04	0.01	–0.002		–0.02		–0.02	0.02
	(0.09)	(0.09)	(0.10)		(0.10)		(0.08)	(0.08)
ILS	0.07	0.10	1.30	*	1.09	*	0.46	0.47
	(0.53)	(0.51)	(0.57)		(0.55)		(0.42)	(0.42)
Number of children								
OLS	0.05	0.05	0.04		0.02		–0.03	0.02
	(0.06)	(0.06)	(0.06)		(0.06)		(0.05)	(0.05)
ILS	0.04	0.05	0.65	*	0.55	*	0.23	0.24
	(0.26)	(0.26)	(0.28)		(0.27)		(0.21)	(0.21)

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. The six subjective scales range from 1 to 5. Larger numbers indicate more positive ratings. All models are estimated with control variables described in Appendix Table A-2. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

4.3 'Motherhood penalty'?

Table 4a presents the fertility effects by ILS and OLS estimation for mothers. Consistent with the estimated Π_1 and θ_1 shown in Table 2, we found that mothers with more children tend to experience significantly greater overall subjective well-being, reporting a premium of 0.99 points when having more than one child, with an increase of 0.46 points for each additional child. However, we found no significant fertility effects for either the pair of time-use outcomes or the logged personal income last month, again, with the exception of hours taking care of family members for number of children marginally significant. All three pairs of coefficients are negative. As mentioned earlier, while the negative coefficients for working hours and income are intuitive, the negative coefficients for hours taking care of family members could be due to the fact that this measure is rather imprecise. These ILS results do not directly speak to household specialisation, although the association of fertility and subjective well-being suggests that greater involvement in childrearing is satisfying to mothers.

The OLS counterparts for fertility effects on subjective well-being are both insignificant. For example, when having more than one child, the coefficient on fertility is –0.03 and is insignificant based on OLS estimates, while the ILS coefficient is 0.99 and is significant; with each additional child, while the OLS coefficient is insignificant and as small as 0.02, the ILS coefficient is significant and is 0.46. Note that the coefficients on fertility for time spent working and caring for family members, though insignificant or only marginally significant in both the ILS and the OLS estimates, have the opposite signs in the two models. These comparisons between results using the two approaches show that the regular OLS regressions are likely to result in biased estimates and that the IV approach may help rectify them.

Table 4b presents details of the link between fertility and the six subjective well-being outcomes for mothers. We found that, with more children, mothers tend to be happier and more satisfied with life, with premiums both of 1.31 points in self-rated happiness and in life satisfaction with more than one child, and with premiums both of 0.61 with each additional child. They also report marginally significant premiums in self-rated social ability: the coefficient on fertility is 0.88 with more than one child being the endogenous variable and is 0.41 with each additional child being the endogenous variable. For mothers, unlike for fathers, the effects of fertility for both of the two self-confidence variables are insignificant. This gender difference is suggestive of household specialisation. Mothers, who bear a disproportionate responsibility for family care relative to fathers, may be more likely to reap premiums in happiness and life satisfaction from the childrearing process, while fathers, who bear more obligations to provide materially for the family relative to mothers, may find that additional children engender a stronger sense of confidence about their careers and their futures.

Table 4a: OLS and ILS estimates of outcome models, mother sample

Dependent variables	Mother (*N*=868)							
	Time use outcomes				**Income outcome**		**Subjective outcome**	
	Logged hours worked per month in 2009		Logged hours taking care of family members last month		Logged personal income last month		Overall subjective scale	
Estimation methods								
More than one child								
OLS	0.09		0.13		–0.37		–0.03	
	(0.09)		(0.39)		(0.48)		(0.07)	
ILS	–0.12		–3.55		–1.50		0.99	*
	(0.51)		(2.19)		(2.70)		(0.42)	
Number of children								
OLS	0.06		0.03		0.02		0.02	
	(0.06)		(0.24)		(0.29)		(0.04)	
ILS	–0.05		–1.65	†	–0.70		0.46	*
	(0.23)		(0.99)		(1.25)		(0.18)	

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Overall subjective scale is the average of six subjective scales ranging from 1 to 5 on overall happiness, life satisfaction, self-confidence in career, self-confidence in the future, quality of social relationship, and social ability. Larger numbers indicate more positive ratings. All models are estimated with control variables described in Appendix Table A-2. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

Table 4b: OLS and ILS estimates of subjective outcome models, mother sample

	Mother (*N*=868)					
	General		**Confidence**		**Social**	
Dependent variables	Self-rated happiness	Life satisfaction	Self-confidence in career	Self-confidence in the future	Self-rated quality of social relationship	Self-rated social ability
Estimation methods						
More than one child						
OLS	0.05	-0.10	-0.12	0.02	-0.02	-0.02
	(0.11)	(0.11)	(0.11)	(0.11)	(0.09)	(0.09)
ILS	1.31 *	1.31 *	0.72	0.98	0.72	0.88 †
	(0.62)	(0.65)	(0.64)	(0.64)	(0.49)	(0.49)
Number of children						
OLS	0.04	-0.05	0.05	0.04	0.04	0.02
	(0.06)	(0.07)	(0.07)	(0.07)	(0.05)	(0.05)
ILS	0.61 *	0.61 *	0.34	0.45	0.33	0.41 †
	(0.27)	(0.29)	(0.29)	(0.29)	(0.23)	(0.22)

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. The six subjective scales range from 1 to 5 . Larger numbers indicate more positive ratings. All models are estimated with control variables described in Appendix Table A-2. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

5. Conclusions and discussion

This analysis contributes theoretically and methodologically to research on 'motherhood penalty' and 'fatherhood premium' in labour force outcomes and related research on subjective well-being. Using a nationally representative dataset from the 2010 CFPS, we examine these topics for contemporary China, which has been buffeted by rapid and tremendous social changes – one of which is weakened norms concerning the gendered division of household labour (Bian, Logan, and Shu 2000; Whyte and Parish 1984; Wolf 1984; Yu and Xie 2011; Zuo and Bian 2001). This analysis provides new evidence concerning the causal and gender-specific effects of fertility on parents' time use, income, and subjective well-being by exploiting the differential implementations of the 'one-child policy' as an IV in estimations.

While we find no fertility effects on income, we find significant effects both on time use and subjective well-being outcomes. With more children, fathers tend to work longer hours, spend less time taking care of family members and report greater subjective well-being. Having more children does not seem to affect mothers objectively in terms of either time use or income, but does lead to significantly better subjective well-being. Among the components of subjective well-being, fathers with more children show greater self-confidence concerning both their careers and the future, while mothers in bigger families report both better overall subjective well-being and better social abilities.

In short, our IV estimations of the causal effects of fertility show premiums for both fathers and mothers and penalties for neither – findings that do not directly support the theory of gendered household specialisation. However, some of the differential effects of fertility on specific domains for mothers versus fathers are consistent with household specialisation. Specifically, compared to mothers, fathers work longer hours and care for their families for fewer hours in response to having more children, which seems a clear indication that fertility leads to greater specialisation in household activities. Compared to fathers, mothers are more likely to reap premiums in happiness, life satisfaction, and social ability from greater fertility, which suggests that mothers derive relatively greater satisfaction from childrearing than do fathers. And finally, the finding that having more children leads to greater career and future confidence for fathers while not for mothers may reflect a tendency for fathers to strengthen their engagement with the labour market in response to a growing family.

To recapitulate, using an innovative method for contemporary rural China, our paper shows overall positive effects of additional fertility for parental outcomes, especially subjective well-being. We interpret this conclusion within the context of China's highly restrictive one-child family planning policy, which generally limits parents to having only one child. Our research suggests that all Chinese parents may do better now that the one-child policy has been abolished and they have been allowed, since October 2015, to have more children (Zhai, Zhang, and Jin 2014). Still, this previously restrictive policy background is a unique setting and affords us the methodological opportunity to study the true causal effects of fertility on parents. We do not wish to generalise our findings to other settings, but we do welcome further research with alternative research designs to address similar issues in other social contexts.

6. Acknowledgements

Early versions of this paper were presented at the 2012 International Sociological Association Research Committee 28 on Social Stratification and Mobility spring meeting in Hong Kong, the 2014 Population Association of America annual meeting in Boston, MA, and the Workshop on Domestic Labour and Fertility Choice in East Asia at the University of Oxford. We are grateful to conference participants, members of the Quantitative Methodology Program at the University of Michigan, two anonymous reviewers, as well as Jennifer Barber, Mary Corcoran, Baochang Gu, Ekaterina Hertog, Man-Yee Kan, Qing Lai, Arland Thornton, Yuying Tong, Xiaogang Wu, and Xiang Zhou for their comments and advice. We also thank N.E. Barr and Cindy Glovinsky for their editorial help.

References

Aassve, A., Goisis, A., and Sironi, M. (2011). Happiness and childbearing across Europe. *Social Indicator Research* 108(1): 65–86. doi:10.1007/s11205-011-9866-x.

Angrist, J.D. and Evans, W.N. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review* 88(3): 450–477.

Angrist, J.D., Imbens, G.W., and Rubin, D.B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434): 444–455. doi:10.1080/01621459.1996.10476902.

Arrow, K.J. (1972). Models of job discrimination. In: Pascal, A.H. (ed.). *Racial discrimination in economic life*. Lexington: D.C. Heath: 83–102.

Arrow, K.J. (1973). The theory of discrimination. In: Ashenfelter, O. and Rees, A. (eds.). *Discrimination in labor markets.* Princeton: Princeton University Press: 3–33.

Baranowska, A. and Matysiak, A. (2011). Does parenthood increase happiness? Evidence for Poland. *Vienna Yearbook of Population Research* 9: 307–325. doi:10.1553/populationyearbook2011s307.

Becker, G.S. (1957). *The economics of discrimination*. Chicago: University of Chicago Press.

Becker, G.S. (1981). *A treatise on the family*. Cambridge: Harvard University Press.

Becker, G.S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics* 3(1): S33–58. doi:10.1086/298075.

Bian, Y., Logan, J.R., and Shu, X. (2000). Wage and job inequalities in the working careers of men and women in Tianjin. In: Entwisle, B. and Henderson, G.E. (eds.). *Re-drawing boundaries: Work, households, and gender in China*. Berkeley: University of California Press: 111–133.

Bianchi, S.M. (1994). Changing economic roles of men and women. In: Farley, R. (ed.). *State of the union: Social trends*. New York: Russell Sage: 107–154.

Billari, F.C. and Kohler, H. (2009). Fertility and happiness in the XXI century: Institutions, preferences and their interactions. Paper presented at XXVI IUSSP International Population Conference, Marrakesh, Morocco, September 27 to October 2, 2009.

Blank, R. (1990). Are part-time jobs bad jobs? In: Burtless, G. (ed.). *A future of lousy jobs? The changing structure of U.S. wages*. Washington, D.C.: Brookings Institution: 123–155.

Blau, F.D. (2012). *Gender, inequality, and wages*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199665853.001.0001.

Budig, M.J. and England, P. (2001). The wage penalty for motherhood. *American Sociological Review* 66(2): 204–225. doi:10.2307/2657415.

China Administration Council (1950). Marriage law of the People’s Republic of China. People’s Republic of China Laws and Regulations: Item 5.

China State Ethnic Affairs Commission (1999). White paper of policies on China’s ethnic minorities and the practice (Zhong guo de shao shu min zu zheng ce ji qi shi jian bai pi shu) (In Chinese).

Chu, J. (2001). Prenatal sex determination and sex-selective abortion in rural central China. *Population and Development Review* 27(2): 259–281. doi:10.1111/j.1728-4457.2001.00259.x.

Corcoran, M.E. and Courant, P.N. (1987). Sex-role socialization and occupational segregation: An exploratory investigation. *Journal of Post Keynesian Economics* 9(3): 330–346. doi:10.1080/01603477.1987.11489627.

Eibach, R.P. and Mock, S.E. (2011). Idealizing parenthood to rationalize parental investments. *Psychological Science* 22(2): 203–208. doi:10.1177/0956797610397057.

England, P. (1992). *Comparable worth: Theories and evidence*. Hawthorne: Aldine de Gruyter.

Filer, R. (1985). Male–female wage differences: The importance of compensating differentials. *Industrial and Labor Relations Review* 38(3): 426–437. doi:10.1177/001979398503800309.

Glauber, R. (2007). Marriage and the motherhood wage penalty among African Americans, Hispanics, and Whites. *Journal of Marriage and Family* 69(4): 951–961. doi:10.1111/j.1741-3737.2007.00423.x.

Glauber, R. (2008). Race and gender in families and at work: The fatherhood wage premium. *Gender & Society* 22(1): 8–30. doi:10.1177/0891243207311593.

Goldin, C. (1995). Career and family: College women look to the past. Cambridge: National Bureau of Economic Research (Working paper No. 5188).

Goldscheider, F., Bernhardt, E., and Lappegård, T. (2015). The gender revolution: A framework for understanding changing family and demographic behavior. *Population and Development Review* 41(2): 207–239. doi:10.1111/j.1728-4457.2015.00045.x.

Gough, M. and Noonan, M. (2013). A review of the motherhood wage penalty in the United States. *Sociology Compass* 7(4): 328–342. doi:10.1111/soc4.12031.

Greenhalgh, S. (2008). *Just one child: Science and policy in Deng's China*. Berkeley: University of California Press. doi:10.1525/california/9780520253384.001.0001.

Gronau, R. (1988). Sex-related wage differentials and women's interrupted careers: The chicken or the egg. *Journal of Labor Economics* 6(3): 277–301. doi:10.1086/298184.

Gu, B., Wang, F., Guo, Z., and Zhang, E. (2007). China's local and national fertility policies at the end of the twentieth century. *Population and Development Review* 33(1): 129–147. doi:10.1111/j.1728-4457.2007.00161.x.

Guo, Z., Liu, J., and Song, J. (2002). China's family planning policy and family structure in the future (Wo guo xian xing sheng yu zheng ce yu wei lai de jia ting jie gou). (In Chinese.) *Chinese Journal of Population Science* (Zhong guo ren kou ke xue) 1: 1–11.

Hannum, E. (2005). Market transition, educational disparities, and family strategies in rural China: New evidence on gender stratification and development. *Demography* 42(2): 275–299. doi:10.1353/dem.2005.0014.

Harkness, S. and Waldfogel, J. (2003). The family gap in pay: Evidence from seven industrialized countries. *Research in Labor Economics* 22: 369–413. doi:10.1016/S0147-9121(03)22012-4.

Hauser, S. and Xie, Y. (2005). Temporal and regional variation in earnings inequality: Urban China in transition between 1988 and 1995. *Social Science Research* 34(1): 44–79. doi:10.1016/j.ssresearch.2003.12.002.

Hill, M. (1979). The wage effects of marital status and children. *Journal of Human Resources* 14(4): 579–594. doi:10.2307/145325.

Hochschild, A. and Machung, A. (1989). *The second shift: Working parents and the revolution at home*. New York: Viking.

Hoffman, L.W. and Hoffman, M.L. (1973). The value of children to parents. In: Fawcett, J.T. (ed.). *Psychological perspectives on population.* New York: Basic Books: 19–76.

Hoffman, L.W., Thornton, A., and Manis, J.D. (1978). The value of children to parents in the United States. *Journal of Population* 1(2): 91–131. doi:10.1007/BF01277597.

Jacobsen, J., Pearce III, J.W., and Rosenbloom, J. (1999). The effects of child-bearing on married women's labor supply and earnings: Using twin births as a natural experiment. *Journal of Human Resources* 34(3): 449–474. doi:10.2307/146376.

Joshi, H. and Newell, M. (1989). *Pay differentials and parenthood: Analysis of men and women born in 1946*. Coventry: University of Warwick Institute for Employment Research.

Killewald, A. (2013). A reconsideration of the fatherhood premium: Marriage, residence, biology, and the wages of fathers. *American Sociological Review* 78(1): 96–116. doi:10.1177/0003122412469204.

Killewald, A. and Gough, M. (2013). Does specialization explain marriage penalties and premiums? *American Sociological Review* 78(3): 477–502. doi:10.1177/0003122413484151.

Kohler, H., Behrman, J.R., and Skytthe, A. (2005). Partner + children = happiness? The effect of partnership and fertility on well-being. *Population and Development Review* 31(3): 407–445. doi:10.1111/j.1728-4457.2005.00078.x.

Korenman, S. and Neumark, D. (1992). Marriage, motherhood, and wages. *Journal of Human Resources* 27(2): 233–255. doi:10.2307/145734.

Kravdal, Ø. (2013). Reflections on the search for fertility effects on happiness. Oslo: Department of Economics, University of Oslo (working paper 10/2013).

Lavely, W., Xiao, Z., Li, B., and Freedman, R. (1990). The rise in female education in China: National and regional patterns. *The China Quarterly* 121: 61–93. doi:10.1017/S0305741000013515.

Loh, E.S. (1996). Productivity differences and the marriage premium for white males. *Journal of Human Resources* 31(3): 566–589. doi:10.2307/146266.

Lundberg, S. and Rose, E. (2000). Parenthood and the earnings of married men and women. *Labor Economics* 7(6): 689–710. doi:10.1016/S0927-5371(00)00020-8.

Margolis, R. and Myrskylä, M. (2011). A global perspective on happiness and fertility. *Population and Development Review* 37(1): 29–56. doi:10.1111/j.1728-4457.2011.00389.x.

Maurer-Fazio, M., Rawski, T.G., and Zhang, W. (1999). Inequality in the rewards for holding up half the sky: Gender wage gaps in China's urban labour market, 1988–1994. *The China Journal* 41: 55–88. doi:10.2307/2667587.

Meisner, M. (1999). *Mao's China and after: A history of the People's Republic*. New York: Free Press.

Meng, X. (1993). Determination and discrimination: Female wages in China's rural TVP industries. *Economics Division Working Papers* 93/1: 20–22.

Miller, A.R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics* 24(3): 1071–1100. doi:10.1007/s00148-009-0296-x.

Nelson, S.K., Kushlev, K., English, T., Dunn, E.W., and Lyubomirsky, S. (2013). In defense of parenthood: Children are associated with more joy than misery. *Psychological Science* 24(1): 3–10. doi:10.1177/0956797612447798.

Neumark, D. and Korenman, S. (1994). Sources of bias in women's wage equations: Results using sibling data. *Journal of Human Resources* 29(2): 379–405. doi:10.2307/146103.

Noonan, M. (2001). The impact of domestic work on men's and women's wages. *Journal of Marriage and Family* 63(4): 1134–1145. doi:10.1111/j.1741-3737.2001.01134.x.

Noonan, M. and Corcoran, M. (2004). The mommy track and partnership: Temporary delay or dead end? *Annals of the American Academy of Sociology and Political Science* 596(1): 130–150. doi:10.1177/0002716204268773.

Oppenheimer, V.K. (1997). Women's employment and the gain to marriage: The specialization and trading model. *Annual Review of Sociology* 23: 431–453. doi:10.1146/annurev.soc.23.1.431.

Parish, W. (1981). Egalitarianism in Chinese society. *Problems of Communism* 29: 37–53.

Parish, W.L. and Busse, S. (1998). Gender and work. In: Tang, W. and Parish, W.L. (eds.). *Market transition in urban China: The changing social contract.* Cambridge: Cambridge University Press: 14.

Peng, P. (1997). *China family planning encyclopedia (Zhong guo ji hua sheng yu quan shu)*. Beijing: China Population Press (Zhong guo ren kou chu ban she) (In Chinese).

Phelps, E.S. (1972). The statistical theory of racism and sexism. *American Economic Review* 62: 659–661.

Polachek, S. (1985). Occupational segregation: A defense of human capital predictions, and Reply to England. *Journal of Human Resources* 20(3): 437–440, 444. doi:10.2307/145892.

Schultz, P.T. (1981). *Economics of population*. Reading: Addison-Wesley.

Smock, P.J., Manning, W.D., and Gupta, S. (1999). The effect of marriage and divorce on women's economic well-being. *American Sociological Review* 64(6): 794–812. doi:10.2307/2657403.

Song, L. (2009). The effect of the Cultural Revolution on educational homogamy in urban China. *Social Forces* 88(1): 257–270. doi:10.1353/sof.0.0246.

Thornton, A. and Lin, H-S. (1994). *Social change and the family in Taiwan*. Chicago: University of Chicago Press.

Treiman, D.J. (2013). Trends in educational attainment in China. *Chinese Sociological Review* 45(3): 3–25. doi:10.2753/CSA2162-0555450301.

Waite, L.J. and Gallagher, M. (2000). *The case for marriage: Why married people are happier, healthier, and better off financially.* New York: Broadway Books.

Waldfogel, J. (1997). The effects of children on women's wages. *American Sociological Review* 62(2): 209–217. doi:10.2307/2657300.

Waldfogel, J. (1998a). The family gap for young women in the United States and Britain: Can maternity leave make a difference? *Journal of Labor Economics* 16(3): 505–545. doi:10.1086/209897.

Waldfogel, J. (1998b). Understanding the 'family gap' in pay for women with children. *Journal of Economic Perspectives* 12(1): 137–156. doi:10.1257/jep.12.1.137.

Wang, W., Jing, H., and Zhang, P. (2013). Has China's family planning policy lowered individuals' level of happiness? (Ji hua sheng yu zheng ce jiang di le ju min de xing fu gan ma?) *Population Research* (*Ren kou yan jiu*) 37: 102–112.

White, M.P. and Dolan, P. (2009). Accounting for the richness of daily activities. *Psychological Science* 20(8): 1000–1008. doi:10.1111/j.1467-9280.2009.02392.x.

Whyte, M.K. (2003). *China's revolutions and intergenerational relations*. Ann Arbor: Center for Chinese Studies, University of Michigan. doi:10.3998/mpub.19840.

Whyte, M.K. (2010). *Myth of the social volcano: Perceptions of inequality and distributive injustice in contemporary China*. Stanford: Stanford University Press.

Whyte, M.K. and Parish, M.L. (1984). *Urban life in contemporary China*. Chicago: University of Chicago Press.

Wolf, M. (1984). Marriage, family, and the state in contemporary China. *Pacific Affairs* 57(2): 213–236. doi:10.2307/2759124.

Wu, X. and Song, X. (2010). Gender inequality in education and employment: China's urban labor markets in transition, 1982–2005. Paper presented at annual meeting of the Population Association of America, Dallas, TX, April 15 to 17, 2010.

Wu, X. and Zhang, Z. (2010). Changes in educational inequality in China, 1990–2005: Evidence from the population census data. *Research in Sociology of Education* 17: 123–152. doi:10.1108/S1479-3539(2010)0000017007.

Xie, Y. (2011). Evidence-based research on China: A historical imperative. *Chinese Sociological Review* 44(1): 14–25. doi:10.2753/csa2162-0555440103.

Xie, Y. (2013). Gender and family in contemporary China. Research Report No. 13–808. Ann Arbor, MI: Population Studies Center, University of Michigan.

Xie, Y. and Hannum, E. (1996). Regional variation in earnings inequality in reform-era urban China. *American Journal of Sociology* 101(4): 950–992. doi:10.1086/230785.

Xie, Y., Zhang, X., Li, J., Yu, X., and Ren, Q. (2013). *Wellbeing development report of China 2013 (Zhong guo min sheng fa zhan bao gao 2013)*. (In Chinese.) Beijing: Peking University Press.

Xu, X., Ji, J., and Tung,Y-Y. (2000). Social and political assortative mating in urban China. *Journal of Family Issues* 21(1): 47–77. doi:10.1177/019251300021001003.

Yu, J. and Xie, Y. (2011). The varying display of 'gender display': A comparative study of mainland China and Taiwan. *Chinese Sociological Review* 44(2): 5–30. doi:10.2753/CSA2162-0555440201.

Yu, J. and Xie, Y. (2013). Social change and trends in determinants of entry to first marriage (She hui bian qian yu chu hun ying xiang yin su de bian hua). (In Chinese.) *Sociological Research* (*She Hui Xue Yan Jiu*).

Zhai, Z., Zhang, X., and Jin, Y. (2014). Demographic consequences of an immediate transition to a universal two-child policy (Li ji quan mian fang kai er tai zheng ce de ren kou xue hou guo fen xi). (In Chinese.) *Population Research* (*Ren kou yan jiu*) 2: 3–17.

Zhang, Y., Hannum, E., and Wang, M. (2008). Gender-based employment and income differences in urban China: Considering the contributions of marriage and parenthood. *Social Forces* 86(4): 1529–1560. doi:10.1353/sof.0.0035.

Zuo, J. and Bian, Y. (2001). Gendered resources, division of housework, and perceived fairness: A case in urban China. *Journal of Marriage and Family* 63: 1122–1133. doi:10.1111/j.1741-3737.2001.01122.x.

Appendix

Table A-1: Exemptions policy to have a second child with one girl

Condition	Province
The parents live in mountain area, rural residents, one girl only	Beijing; Tianjin; Shanxi; Inner Mongol; Jilin; Heilongjiang; Zhejiang; Anhui; Fujian; Jiangxi; Henan; Hubei; Hunan; Guangdong; Chongqing; Guizhou; Shannxi; Gansu.
The parents work in mining industry and directly work in mines, one girl only	Hebei; Jiangsu; Zhengjiang; Anhui; Shandong; Henan.
Mother rural, one girl only	Guangxi.
Mother rural, one girl only and with rural registration	Liaoning; Shandong.
Mother rural, one girl only, father living with his parents-in-law, mother without brothers	Jiangsu.
Mother rural, one girl only, father without brothers and with only one sister	Jiangsu.
Mother rural, one girl only, spouse living in coastal farming areas	Jiangsu.
Mother rural, one girl only, one of the parents in marine fishing	Jiangsu.
Both parents rural, one of the parents having non-heritable physical disability, one girl only	Jiangsu.
One of the parents works as contract worker in farming industry, one girl only	Jilin.
One of the parents works in marine fishing industry, one girl only	Shandong.
One of the parents has non-heritable physical disability, one girl only	Shandong.

Source: Population and Family Planning Commission of Shanxi website. http://www.sxrk.gov.cn/Article.jsp?ArticleID=4623

Table A-2: Differences in means for control variables

	Difference in means by gender of the first child (ref.=male)			
Variables	Father (*N*=1,123)		Mother (*N*=868)	
Agriculture	–0.016		0.020	
	(0.030)		(0.033)	
Migrant	0.018		0.001	
	(0.018)		(0.021)	
Age	0.423		–0.225	
	(0.353)		(0.385)	
Age at first birth	0.508	*	0.043	
	(0.196)		(0.205)	
Illiterate or semi illiterate	0.040	†	0.029	
	(0.020)		(0.030)	
Primary	–0.058	*	–0.002	
	(0.027)		(0.032)	
Junior middle	0.023		–0.028	
	(0.030)		(0.031)	
Senior middle	–0.007		0.003	
	(0.018)		(0.017)	
Associate college or above	0.002		–0.002	
	(0.010)		(0.011)	
Age gap between the oldest child and the youngest child	1.253	***	1.390	***
	(0.209)		(0.240)	
Youngest child coresidence	0.007		0.002	
	(0.008)		(0.011)	
Spouse coresidence	0.004		0.002	
	(0.004)		(0.009)	
Grandfather not coresiding	0.023		0.020	
	(0.028)		(0.013)	
Grandfather coresiding, under age 70	–0.022		–0.025	*
	(0.027)		(0.012)	
Grandfather coresiding, age 70 or above	–0.001		0.005	
	(0.014)		(0.005)	
Grandmother not coresiding	0.002		0.013	
	(0.029)		(0.014)	
Grandmother coresiding, under age 70	–0.030		–0.018	
	(0.028)		(0.013)	
Grandmother coresiding, age 70 or above	0.028	†	0.005	
	(0.015)		(0.005)	

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Province of the respondent's residential registration is also controlled for as a set of dummy variables to single out the regional fixed effects. P-value of Pearson's Chi-square test for association between gender of the first child and province of residential registration is 0.514 for fathers, and 0.764 for mothers. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.

Table A-3a: OLS and Wald estimates of outcome models, father sample

Dependent variables	Father (N=1,123)							
	Time use Outcomes				**Income outcome**		**Subjective outcome**	
	Logged hours worked per month in 2009		Logged hours taking care of family members last month		Logged personal income last month		Overall subjective scale	
Estimation methods								
More than one child								
OLS	–0.07		–0.08		–1.26	***	–0.08	*
	(0.07)		(0.24)		(0.32)		(0.04)	
Wald	0.54	*	–0.52		1.34		0.22	
	(0.21)		(1.12)		(1.52)		(0.19)	
Number of children								
OLS	–0.04		0.02		–0.74	**	–0.05	†
	(0.03)		(0.16)		(0.22)		(0.03)	
Wald	0.32	**	-0.31		0.79		0.13	
	(0.12)		(0.66)		(0.90)		(0.11)	

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Overall subjective scale is the average of six subjective scales ranging from 1 to 5 on overall happiness, life satisfaction, self-confidence in career, self-confidence in the future, quality of social relationship, and social ability. Larger numbers indicate more positive ratings. All models are estimated without any control variables.†p<0.10; *p<0.05; **p<0.01; ***p<0.001.

Table A-3b: OLS and Wald estimates of outcome models, mother sample

Dependent variables	Mother (*N*=868) Time use outcomes: Logged hours worked per month in 2009		Time use outcomes: Logged hours taking care of family members last month		Income outcome: Logged personal income last month		Subjective outcome: Overall subjective scale	
Estimation methods								
More than one child								
OLS	–0.07		0.70	*	–2.39	***	–0.18	***
	(0.06)		(0.27)		(0.39)		(0.04)	
Wald	–0.20		0.75		–1.89		0.40	†
	(0.28)		(1.23)		(1.76)		(0.21)	
Number of children								
OLS	–0.04		0.46	*	–1.34	***	–0.11	***
	(0.04)		(0.18)		(0.27)		(0.03)	
Wald	–0.11		0.42		–1.06		0.22	†
	(0.16)		(0.69)		(0.99)		(0.12)	

Note: 2010 CFPS. Standard errors are reported in the parentheses. The sample is restricted to rural registration, Han ethnicity, eligible provinces with rural registration exemption and eligible cases for all the independent variables. Missing values of the outcome variables were imputed using multiple imputations. Overall subjective scale is the average of six subjective scales ranging from 1 to 5 on overall happiness, life satisfaction, self-confidence in career, self-confidence in the future, quality of social relationship, and social ability. Larger numbers indicate more positive ratings. All models are estimated without any control variables. †p<0.10; *p<0.05; **p<0.01; ***p<0.001.