



SCAPE

Singapore Centre for Applied and Policy Economics

**Does the IV estimator establish causality?
Re-examining Chinese fertility-growth
relationship**

by

Tilak Abeysinghe and Gu Jiaying

**Department of Economics
SCAPE Working Paper Series
Paper No. 2009/02 – Sept 2009**

<http://www.fas.nus.edu.sg/ecs/pub/wp-scape/0902.pdf>

Does the IV estimator establish causality? Re-examining Chinese fertility-growth relationship

Tilak Abeysinghe and Gu Jiaying
Department of Economics
National University of Singapore
September 2009

Abstract

The instrumental variable (IV) estimator in a cross-sectional or panel regression model is often taken to provide valid causal inference from contemporaneous correlations. In this exercise we point out that the IV estimator, like the OLS estimator, cannot be used effectively for causal inference without the aid of non-sample information. We present three possible cases (lack of identification, accounting identities, and temporal aggregation) where IV estimates could lead to misleading causal inference. In other words, a non-zero IV estimate does not necessarily indicate a causal effect nor does the causal direction. In this light, we re-examine the relationship between Chinese provincial birth rates and economic growth. This exercise highlights the potential pitfalls of using too much temporal averaging to compile the data for cross sectional and panel regressions and the importance of estimating both (x on y and y on x) regressions to avoid misleading causal inferences. The GMM-SYS results from dynamic panel regressions based on five-year averages show a strong negative relationship running both ways, from births to growth and growth to births. This outcome, however, changes to a more meaningful one-way relationship from births to growth if the panel analysis is carried out with the annual data. Although falling birth rates in China have enhanced the country's growth performance, it is difficult to attribute this effect solely to the one-child policy implemented after 1978.

Key words: IV estimator and causality inference, identification, accounting identities, temporal aggregation, spurious causality, Chinese provincial growth and fertility relationship.

JEL: C23, J13, O53

Authors would like to thank Sam Ouliaris, Chan Kok Hoe and the participants of the SER Conference 2009 for their valuable comments on this exercise.

1. Introduction

Growth regressions based on instrumental variables from cross-sectional data are often used to examine causality between economic growth and a variable of interest.¹ Does the IV estimator always deliver valid causal inference in this context is a question of interest. The IV estimator has come under some scrutiny recently, especially in relation to the omitted variable bias and “natural instruments” or instruments from natural experiments (for both pros and cons see Angrist and Krueger, 2001, Rosenzweig and Wolpin, 2000, Deaton, 2009, Heckman and Urzua, 2009, Imbens 2009 and the references in them).² The objective of this exercise is not to diverge into this literature but to highlight some major concerns of using the IV estimator in growth regressions for causality inference and then to analyze the Chinese fertility-growth relationship.

China’s experience of rapid economic growth and falling fertility stand out as a basket case for assessing the usefulness of IV methods for causal inference. Unlike the developed industrial countries that experienced reduced family size as a choice outcome, China’s fertility decline was precipitated by the one-child policy. This decline coincided with China’s rapid economic growth brought about by the market oriented open economy growth strategies. A priori, therefore, one may argue that causality should run from the policy induced birth rate to economic growth as suggested by Li and Zhang (2007).³ If we go one step further to examine the reverse causality, we could expect the growth rate to have little explanatory power on the policy induced birth rate. Alternatively, as predicted by the price theoretic model of fertility decisions, we could expect a

¹ For example, for trade-growth causality see Frankel and Romer (1999) and finance-growth causality see Levine et al. (2000).

² Deaton (2009) places his skepticism on the IV estimator as: “Economists’ claims to methodological superiority based on instrumental variables ring particularly hollow when it is economists themselves who are often misled. My argument is that for both economists and non-economists, the direct consideration of the reduced form is likely to generate productive lines of enquiry.”

³ Li and Zhang (2007), based on IV estimates, drew a conclusion in support of the neo-Malthusian thesis that high birth rates hamper economic growth and therefore, China’s one-child policy was growth enhancing.

positive effect of higher incomes on fertility if parents try to shun the one-child policy and go for a larger family size.

After highlighting three situations (lack of identification, accounting identities, and temporal aggregation) where the IV estimator could produce misleading causal inference we move on to analyze the Chinese data by estimating a growth regression and birth rate regression with both five-yearly and annual data. The results are summarized in the concluding section.

2. IV estimator and misleading causal effects

For causality inference from the IV estimator a researcher would typically estimate one of the following two regressions in a pure cross sectional or non-dynamic panel setting:

$$\mathbf{y}_1 = \mathbf{y}_2\gamma_1 + \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{u} \quad (1)$$

$$\mathbf{y}_2 = \mathbf{y}_1\gamma_2 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{v} \quad (2)$$

where $\mathbf{y}_1, \mathbf{y}_2, \mathbf{u}, \mathbf{v}$ are $(N \times 1)$, \mathbf{X}_1 is $(N \times K_1)$, \mathbf{X}_2 is $(N \times K_2)$, N is the number of observations and γ s and β s are the parameters. Suspecting endogeneity,⁴ suppose z is used as an instrument for y_2 in (1) or w is used as an instrument for y_1 in (2). Now using partitioned matrices in the standard IV formula we can obtain

$$\hat{\gamma}_{IV} = (\mathbf{z}'\mathbf{M}_1\mathbf{y}_2)^{-1}\mathbf{z}'\mathbf{M}_1\mathbf{y}_1 = (\mathbf{z}^*\mathbf{y}_2^*)^{-1}\mathbf{z}^*\mathbf{y}_1^* = \sum_{i=1}^N z_i^* y_{i1}^* / \sum_{i=1}^N z_i^* y_{i2}^* \quad (3)$$

$$\hat{\gamma}_{2,IV} = (\mathbf{w}'\mathbf{M}_2\mathbf{y}_1)^{-1}\mathbf{w}'\mathbf{M}_2\mathbf{y}_2 = (\mathbf{w}^*\mathbf{y}_1^*)^{-1}\mathbf{w}^*\mathbf{y}_2^* = \sum_{i=1}^N w_i^* y_{i2}^* / \sum_{i=1}^N w_i^* y_{i1}^* \quad (4)$$

where $\mathbf{M}_1 = (\mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1')$ and $\mathbf{M}_2 = (\mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2')$. Based on the symmetry and idempotent properties of these matrices we have written the variables with asterisks (e.g., $\mathbf{z}^* = \mathbf{M}_1\mathbf{z}$, $\mathbf{y}_1^* = \mathbf{M}_1\mathbf{y}_1$) that are purged of the effects of \mathbf{x} variables. The numerators and denominators of (3) and (4) scaled by N converge to the corresponding covariances and if there is *endogeneity* and by the definition of the instruments these covariances are not zero. Given the

⁴ IV estimators are used to correct for OLS bias resulting from endogeneity, measurement errors, and unobserved omitted variables. In this exercise, the focus is only on endogeneity.

interdependence, therefore, the role of the IV estimator reduces to establishing the causal effect, not the causal direction.

Unfortunately there are circumstances where the IV estimator may provide misleading causal effects. We present three such possible cases. First, if $X_1 = X_2 = X$ in (1) and (2) and if the same instrument is used then we get $\hat{\gamma}_{2,IV} = 1 / \hat{\gamma}_{1,IV}$. Even if the instruments are different the same X leads to $\hat{\gamma}_{1,IV} \rightarrow \gamma_1 = 1 / \gamma_2$ and $\hat{\gamma}_{2,IV} \rightarrow \gamma_2 = 1 / \gamma_1$ because one regression is a simple algebraic transformation of the other. This identification problem is not an unlikely scenario when researchers specify only one of the two equations for estimation.

Second, consider a case where endogeneity results from an explicit or implicit accounting or definitional identity, which is quite common in macro models. To illustrate the case consider the following simple macro model, a simplification from a large literature on the debate on export-led growth versus growth-led exports (see Giles and Williams, 2000, for a survey):

$$s_t^x x_t = \gamma_1 y_t + \beta_1 x_t^* + u_t \quad (5a)$$

$$y_t \equiv s_t^x x_t + s_t^n n_t \quad (5b)$$

where $u_t \sim iid(0, \sigma_u^2)$, y_t is the GDP growth rate, x_t and n_t are respectively the growth rates of exports and non-exports in the GDP identity, s_t^x and s_t^n are the corresponding shares in GDP and x_t^* is an index that captures export promotion policies. To assess the causal effect of exports on GDP growth it is common to replace exports and non-exports in the GDP identity with some derivatives of their production functions and arrive at a behavioral equation. A simplified version of this equation is:

$$y_t = \gamma_2 s_t^x x_t + \beta_2 l_t + v_t \quad (5c)$$

where l_t is the population growth rate and it is assumed $v_t \sim iid(0, \sigma_v^2)$.⁵

Given that (5a) and (5c) constitute an exactly identified system of equations we can easily derive the IV estimates. Since our focus is on γ_2 and using the identity (5b) we obtain

⁵ Since Feder (1982) the primary focus of many researchers has been on a variant of (5c) and not (5a).

$$\hat{\gamma}_{2,IV} = \frac{\sum l_t^2 \sum x_t^* y_t - \sum x_t^* l_t \sum l_t y_t}{\sum l_t^2 \sum x_t^* s_t^x x_t - \sum x_t^* l_t \sum l_t s_t^x x_t} \quad (6)$$

$$= 1 + \frac{\sum l_t^2 \sum x_t^* s_t^n n_t - \sum x_t^* l_t \sum l_t s_t^n n_t}{\sum l_t^2 \sum x_t^* s_t^x x_t - \sum x_t^* l_t \sum l_t s_t^x x_t}.$$

This shows that $\hat{\gamma}_{2,IV}$ is a mixture of the identity effect (given by unity) and the causal effect. If there are no spillover effects from the export sector to the non-export sector ($Cov(x_t^*, s_t^n n_t) = 0$), and assuming that the denominator of (6) is positive, $Cov(l_t, s_t^n n_t) > 0$, and $Cov(x_t^*, l_t) \geq 0$, we may get $\hat{\gamma}_{2,IV} \leq 1$. If the causal effect is positive we could expect $\hat{\gamma}_{2,IV} > 1$.⁶ In cross-country empirical studies, those who used OLS on a general form of (5c) obtained $\hat{\gamma}_{2,OLS} < 1$ whereas those who used IV obtained estimates slightly bigger than unity (see for example, McNab and Moore, 1998, Tables 1 and 3). This, however, does not amount to establishing a causal relationship because of the presence of the identity effect.⁷ As emphasized by Hendry (1995, p. 790) an identity does not represent a causal relationship.

Third, a case that would be of particular interest, is one-way causality. For example, if $\gamma_2 = 0$ in (2), by substituting (2) into (4) and noting $\mathbf{M}_2 \mathbf{X}_2 = \mathbf{0}$, we obtain $\hat{\gamma}_{2,IV} \rightarrow 0$. In this case, both OLS and IV estimators are consistent for γ_1 and only the IV estimator is consistent for γ_2 . Therefore, we could estimate both regressions by both OLS and IV and compare the results and then make causal inferences.

This theoretical outcome, however, may not be materialized in practice because of the temporal aggregation problem. It is common in panel regressions to use long-term (five years or more) averages to capture long-term effects. This, however, does not necessarily make causal inference easier. Even in VAR processes with contemporaneously uncorrelated error terms, as temporal

⁶ Monte Carlo results based on a more general model similar to that of Feder (1982) re-affirmed these theoretical results.

⁷ The result in (6) is suggestive of using $\hat{\gamma}_{2,IV} - 1$ to test for causality when an identity problem is involved. We have not investigated this possibility in detail.

aggregation increases, the transformed error terms become highly contemporaneously correlated by absorbing the causal information contained in the lagged variables and make causal inference difficult.⁸ To illustrate the problem in focus in this paper, consider the following data generating process:

$$\begin{pmatrix} 1 & -g_1 \\ -g_2 & 1 \end{pmatrix} \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix} \quad (7)$$

where $u_t \sim iid(0, \sigma_u^2)$, $v_t \sim iid(0, \sigma_v^2)$ and $E(u_t u_s) = 0, \forall t, s$. Now consider estimating one of the following models from temporally aggregated (or averaged) data which are represented by the upper case letters:

$$Y_\tau = \alpha_1 + \gamma_1 X_\tau + \beta_1 Y_{\tau-1} + U_\tau \quad (8a)$$

$$X_\tau = \alpha_2 + \gamma_2 Y_\tau + \beta_2 X_{\tau-1} + V_\tau. \quad (8b)$$

Note that parameter values also change with temporal aggregation. To assess the performance of the IV estimator under temporal aggregation we carried out a Monte Carlo simulation by generating data from (7) based on $N(0,1)$ u_t and v_t at the frequency $t = 1, 2, \dots, 6000$ and then averaging y_t and x_t over 12 periods to obtain non-overlapping averages at frequency $\tau = 1, 2, \dots, 500$. Then using $X_{\tau-1}$ and $X_{\tau-2}$ as instruments in (8a) and $Y_{\tau-1}$ and $Y_{\tau-2}$ as instruments in (8b) and averaging over 1000 replications we obtained the IV estimates. For brevity we report only a few representative cases in Table 1. We also provide OLS estimates in Table 1 simply for comparison. Although they are not consistent because of the correlation between U_τ and V_τ , they are informative.

⁸ See Rajaguru and Abeysinghe (2002; also references therein) for extensive analytical and Monte Carlo results on how stationary VAR processes shrink towards VAR(0) as temporal aggregation increases and create contemporaneous correlations that take positive, negative or zero values depending on the relative magnitudes of the VAR parameters.

Table 1. OLS and IV estimates from temporally aggregated data

Original parameters				Parameter values after temporal aggregation			
g_1	g_2	b_1	b_2	$\hat{\gamma}_{1,OLS}$	$\hat{\gamma}_{2,OLS}$	$\hat{\gamma}_{1,IV}$	$\hat{\gamma}_{2,IV}$
0.5	0	0.3	0.8	0.70	1.13	0.74	1.22
-0.5	0	0.3	0.8	-0.70	-1.13	-0.75	-1.23
-0.5	0.1	0.3	0.8	-0.65	-1.05	-0.75	-1.31
-0.1	0.5	0.8	0.3	1.06	0.64	1.27	0.74
-0.5	0.5	0.1	0.3	-0.09	-0.12	-0.15	-0.25
-0.5	0.5	0.3	0.1	0.12	0.09	0.20	0.24

The results in Table 1 clearly show that even when $g_2 = 0$ (one-way causality from x to y) $\hat{\gamma}_{2,IV} \neq 0$. The sign of $\hat{\gamma}_2$ is determined by the sign of g_1 in the disaggregate process. An interesting case emerges when both g_1 and g_2 are not zero. Although $\gamma_2 \neq 0$ in this case, the sign of $\hat{\gamma}_{2,IV}$ may not necessarily be that of g_2 ; the sign depends on the sign of the dominant effect between the two-way relationship. Moreover, as the last two rows of the table highlight, there are also possibilities where both parameter estimates $\hat{\gamma}_{1,IV}$ and $\hat{\gamma}_{2,IV}$ may become so small that we may end up with the wrong inference that the two variables are unrelated.

The non-zero IV estimates presented above resulting from (i) a lack of identification, (ii) an accounting identity, and (iii) temporal aggregation highlight the difficulty involved in making causal inference from an IV estimate alone. Although the IV estimators are consistent they lack any causal meaning in these cases. There is also a circularity of the argument that the IV estimator can be used to test causality; we start assuming endogeneity and this is exactly what the IV estimator provides if the testing is done both ways. In general, the IV estimator, like the OLS estimator, cannot be used for causal inference without additional information. Note that the use of natural instruments does not solve these problems.

Making causal inference from contemporaneous correlations is a challenging one. With reference to the temporal aggregation problem, Rajaguru and Abeysinghe (2008) present a solution within a cointegrating framework for pure time series models. If we want to use the IV estimator for

causal inference in a cross sectional setting, then we have to estimate both the regressions, preferably within a well specified structural framework. However, as we shall see in the next section, causal inference from the IV estimator may still need a fare amount of non-sample information.

3. Chinese fertility-growth causality

As discussed in the previous section we estimate two regressions, a growth regression and a birth rate regression. The panel growth regression is essentially that of Li and Zhang (2007):

$$\ln(y_{it} / y_{i,t-1}) = \gamma_1 BR_{it} + \beta_1 \ln y_{i,t-1} + \mathbf{x}'_{it} \boldsymbol{\beta}_1^* + \eta_i + u_{it} \quad (9a)$$

where y is per capita income, BR is the birth rate as conventionally defined, \mathbf{x}_1 is a vector of control variables, and η_i is time-invariant regional effects. Before we move on to the rest of the discussion it is important to understand the exact meaning of the regression (9a). Fig 1 provides a scatter plot of per capita income growth and birth rates by province for time series data over 1954-2002. As we can see, there is virtually no relationship between the two variables. Li and Zhang (2007), however, find a statistically significant negative relationship between the two variables. This is because (9a) is an equivalent transformation of the model

$$\ln y_{it} = \gamma_1 BR_{it} + \beta_1^* \ln y_{i,t-1} + \mathbf{x}'_{it} \boldsymbol{\beta} + \eta_i + u_{it} \quad (9b)$$

with $\beta_1 = \beta_1^* - 1$ which is negative if β_1^* is a fraction. Fig 2 is a scatter plot of the relationship between log per capita income and the birth rate by province over 1953-2002. Now there is a clear negative relationship between the two variables. This negative relationship is exactly what the Li-Zhang regression picks up and therefore the regression (9a) should be interpreted as an income-birth model rather than a growth-birth model.⁹ Nevertheless, we continue to refer to it as a growth-birth model.

Insert Fig 1 and Fig 2

⁹ These type of growth regressions typically involve an initial level of per capita income ($\ln y_{i,t-1}$) which yields a negative estimated coefficient. This is a technical effect because of the equivalence of the transformation involved.

To specify the birth rate regression we follow the literature on price theoretic and relative income approaches to modeling fertility decisions (Easterlin, 1983, Abeyasinghe, 1993):

$$BR_{it} = \gamma_2 \ln y_{it} + \gamma_3 \ln y_{i,t-1} + \beta_2 BR_{i,t-1} + \mathbf{x}'_{2it} \boldsymbol{\beta}_2^* + \mu_i + v_{it} \quad (10a)$$

$$BR_{it} = \gamma_2 \ln(y_{it} / y_{i,t-1}) + (\gamma_2 + \gamma_3) \ln y_{i,t-1} + \beta_2 BR_{i,t-1} + \mathbf{x}'_{2it} \boldsymbol{\beta}_2^* + \mu_i + v_{it} \cdot \quad (10b)$$

In (10) the variables that capture the time cost of women are absorbed into the vector \mathbf{x}_2 and μ_i represent time-invariant regional effects. Both the current income and lagged income ($\ln y_{it}$ and $\ln y_{i,t-1}$) appear in (10a) to capture the relative income effect; an increase in current income relative to a reference income level is expected to increase fertility.¹⁰ Therefore, we expect, $\gamma_2 > 0$ and $\gamma_3 < 0$, and if these two coefficients are of equal magnitudes then the coefficient of $\ln y_{i,t-1}$ in (10b) becomes zero. Obviously the choice of control variables in regressions (9) and (10) is constrained by the availability of the required data.

We present two sets of estimation results for (9) and (10). In the first set, for the purpose of comparison with Li and Zhang (2007), we use data averaged over five year intervals (except for the initial levels) over 1978-2002 for 28 provinces.¹¹ In the second set we use annual data over 1953-2002. For the estimation, we adopt the GMM-SYS methodology that Li and Zhang have adopted that involves estimating the equations in first differences to remove the time-invariant provincial fixed effects and then using a system GMM to estimate growth and BR regressions separately. As for instruments we use the minority proportion in the first set of regressions as in

¹⁰ It is worth mentioning that the non-Malthusian outcome of falling family size as income rises has been a puzzle for economists for long time (Becker and Lewis, 1973; Easterlin, 1983). When income elasticity from fertility regressions continued to produce negative numbers Leibenstein (1975) asked “Are children inferior goods”. This puzzle does not seem to arise within a relative income framework (Abeyasinghe, 1993).

¹¹ The main innovative feature of the Li and Zhang (2007) study was the use the proportion of minority population in each province as an instrument for the birth rate in their growth rate regression. Unfortunately the data on minority proportions are available only at five year intervals. This forced them to use average data over five year time intervals. Ours is an extended sample period. We collected the data from the following sources. Demographic variables like the birth rate and the minority proportion are from Basic Data of China’s Population (1994, 2003). Economic variables are from the Comprehensive Statistical Data and Materials on 55 Years of New China 1949-2004 (2005) and various issues of China Statistical Yearbook (1980-2002) and China Population Statistical Yearbook (1980-2002).

Li and Zhang (2007). All other instruments are the relevant lags of the internal variables (see Bond, 2002; Arellano, 2003; Roodman, 2006).

Table 2 presents the GMM-SYS estimation results based on five-year averaged data over the 1978-2002 period. The results for the growth-birth model (columns 1-3) basically lead to the Li-Zhang conclusion; the reduction in the birth rate has enhanced the growth performance of the Chinese economy. If we examine the reverse causality given by the estimates in the birth-growth model (columns 3-6) we again obtain a highly significant negative effect leading to the conclusion that the higher the growth the lower the birth rate. Does this tell us anything beyond the negative association we observe between the two variables? Since we expected a zero or a positive effect of growth on fertility, these IV estimates need to be doubted. As we have seen in the previous section, it is very likely that too much temporal averaging may have led to this outcome and it would become difficult to make causal inference from either of the two regressions because of the potential distortions reported in Table 2. It should also be noted that China's birth rate started falling well before the implementation of the one-child policy since 1978 (Fig 3); in fact the most precipitous fall occurred during the Cultural Revolution period. Therefore, confining to the sample period after 1978 may entail a sample selection bias that could lead to attributing the positive effect of falling birth rate on growth solely to the one-child policy.

Insert Table 2 and Fig 3

With the annual-data (1953-2002) models we use a number of step dummies to capture some specific effects. The first dummy ($D78_t = 1$ for $t \geq 1978$ and zero otherwise) is introduced to both equations to capture the effect of both the birth control and openness policy reforms since 1978. The second dummy ($D6577_t = 1$ for $1965 \leq t \leq 1977$ and zero otherwise) is introduced to the growth equation to assess the impact of the falling birth rates before the implementation of the one-child policy. The third dummy ($D65_t = 1$ for $t \geq 1965$ and zero otherwise) is introduced to the growth equation to capture the effect of the steady fall in the birth rate since the mid 1960s. In this exercise we observe that the use of individual year dummies to control for the time effect lead to some distortions of the main results. This problem does not arise if we use time dummies for grouped years as in Table 2. We tried regressions with time dummies for three,

four, and five year groups and observed that the main results remain very much un-affected by these year groupings. The three-year grouping automatically accounts for the famine period over 1959-1961 during which the birth rates dived dramatically across the country. Table 3 presents some GMM-SYS regression results based on three-year time dummies. In the table “Famine” refers to the time dummy that stands for the famine period.

In contrast to the results in Table 2, the results in Table 3 and many other regressions we ran with annual data indicate the presence of a clear one-way negative effect running from the birth rate to the growth rate (to per capita income rather). If we start with the birth-growth regression results in Table 3 (columns 4-6) we observe that the growth effect on the BR is rather fragile. In some regressions the growth coefficient becomes positive and significant as predicted by economic theory but the fragility of this effect leads us to conclude that there is no significant growth (income) effect on the birth rate. Unlike the growth (income) effect, the time cost effect, as proxied by the secondary school enrolment, stays persistently negative in all regressions though not necessarily significant all the time. The significant negative coefficient of $D78_t$ captures the level-shift in the birth rate as a result of the one-child policy. The significant positive coefficients of BR_{t-1} and $D78.BR_{t-1}$ show the increase in the autoregressive effect as a result of the birth control policy. Overall, the birth rate dynamics seem to be determined by birth control policies and time cost effects. Although some parents may have shunned the one-child policy and opted for two or more children as their incomes rose this effect is not strong enough to be picked up by the aggregated data.

The results from the growth-birth regressions (columns 1-3) are very instructive. All the control variables have the expected signs and mostly statistically significant. If we focus on the impact of the BR on growth, the regression in column 1 indicates that after 1978 the fall in the BR has enhanced the growth significantly (coefficient of $D78.BR$). We may be tempted to attribute this to the one-child policy. However, as we noted earlier the fertility decline in China started well before the one-child policy. If we introduce the interaction dummy $D6577.BR$ also to the regression (column 2), we notice that there was no change in the slope of the growth-birth relationship, after controlling for the effects of openness on growth, as a result of the one-child policy. In column 3 we replace the above two interaction dummies with $D65.BR$ to pick up the overall negative effect of the BR on growth. What we notice from these results is that there was

no significant relationship between the birth rate and growth before 1965. However, as the birth rate started to fall it has generated a robust income enhancing effect that works out to be about nine percent ($0.016/0.184 = 0.087$) growth in per capita income in the long run for one unit drop in the birth rate. This effect is much lower than that implied by the numbers in Table 2 (about 19%). What we can conclude here is that although falling birth rates have enhanced China's growth as Li and Zhang (2007) have observed we cannot necessarily attribute this to the one-child policy.

Conclusion

Although the IV estimator has become a popular technique in making causal inference from contemporaneous correlations in cross-sectional and panel regressions the results in this exercise highlight some potential pitfalls involved in such analyses. Apart from the common cause problem, lack of identification, accounting identities and temporally aggregated data may render highly invalid inference from IV estimates. However, by estimating both regressions and with the aid of non-sample information we might be able to make causal inference from the IV estimates.

Our analysis of the Chinese fertility-growth relationship based on five-year average data and annual data in a dynamic panel setting reveal the possibility of very different causal inferences. Although it is often thought that it would be better to take long-term averages to assess long-term effects, making causal inference from the resulting contemporaneous correlations may run into difficulties as we have seen in this exercise. The results from the annual data are consistent with our apriori theoretical expectations and therefore are more likely to represent the causal effects between the birth rate and the growth rate. Our extensive analysis leads us to conclude that causality runs from the birth rate to growth (or income) and if there is any growth (income) effect on the birth rate it is negligibly small at a provincial level. The long-run effect generated from the annual data model indicates that a sustained one unit fall in the birth rate has increased China's steady-state per capita income by about nine percent.

References

- Abeysinghe, T. (1993) "Time cost, relative income and fertility in Canada" *Journal of Population Economics*, 6, 189-198.
- Angrist, J.D. and Krueger, A.B. (2001) Instrumental variables and the search for identification: From supply and demand to natural experiments, *Journal of Economic Perspectives*, 15, 69-85.
- Arellano, M. (2003) *Panel Data Econometrics*, Oxford Univ Press, Oxford.
- Becker, G. and Lewis, G. (1973) "On the interaction between the quantity and quality of children" *Journal of Political Economy*, 81, S279-288.
- Bond, S. (2002) "Dynamic panel data models: a guide to micro data methods and practice", *Portuguese Economic Journal*, 1, 141-162.
- Deaton, A. (2009) "Instruments of Development: Randomization in the tropics, and the search for the illusive keys of economic development" NBER Working Paper No: 14690.
- Easterlin, R. A., Pollak, R.A., and Wachter, M.L. (1980) "Toward a more general model of fertility determination: Endogenous preferences and natural fertility" in Easterlin R. A. ed. *Population and Economic Change in Developing countries*. NBER (Chicago: Chicago Univ Press) 81-150.
- Feder, G. (1982) "On exports and economic growth", *Journal of Development Economics*, 12, 59-73.
- Frankel, J.A. and Romer, D. (1999) "Does trade cause growth?", *American Economic Review*, 89, 379-399.
- Giles, J.A. and Williams, C.L. (2000) Export-led growth: a survey of the empirical literature and some non-causality results, Part 1, *Journal of International Trade & Economic Development*, 9, 261-227.
- Heckman, J.J. and Urzua, S. (2009) Comparing IV with structural models: What simple IV can and cannot identify, NBER Working Paper No: 14706.
- Hendry, D. (1995) *Dynamic Econometrics*, Oxford University Press.
- Imbens, G.W. (2009) Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009), Department of Economics, Harvard University.
- Leibenstein, H. (1975) "The economic theory of fertility decline" *Quarterly Journal of Economics*, 89, 1-31.

- Levine, R.; Loayza, N. and Beck, T. (2000) "Financial intermediation and growth: Causality and causes" *Journal of Monetary Economics*, 46, 31-77.
- Li, H. and Zhang, J. (2007) "Do high birth rates hamper economic growth?" *The Review of Economics and Statistics*, 89, 110-117.
- McNab, R.M. and Moore, R. E. (1998) "Trade policy, export expansion, human capital and growth", *Journal of International Trade & Economic Development*, 7, 237-256.
- Rajaguru, G. and Abeysinghe, T. (2002) "The distortionary effects of temporal aggregation on Granger causality", Working Paper wp0204, Department of Economics, National University of Singapore: <http://www.fas.nus.edu.sg/ecs/pub/wp/wp0204.pdf> .
- Rajaguru, G. and Abeysinghe, T. (2008) "Temporal aggregation, cointegration and causality inference", *Economics Letters*, 101, 223-226.
- Roodman, D. (2006) "How to Do xtabond2: An introduction to 'Difference' and 'System' GMM in Stata", Working Paper # 103, Center for Global Development.
- Rosenzweig, M.R. and Wolpin, K.I. (2000) Natural "natural experiments in economics, *Journal of Economic Literature*, 38, 827-874.

Table 2: Five-year average data models

Growth-birth model			Birth-growth model				
Dependent variable: $\Delta \ln y_{it}$			Dependent variable: BR_{it}				
	(1)	(2)	(3)		(4)	(5)	(6)
BR_{it}	-0.026*** (0.007)	-0.030*** (0.007)	-0.029*** (0.007)	$\Delta \ln y_{it}$	-10.841** (5.042)	-10.782** (4.206)	-12.209** (5.77)
$\ln y_{i,t-1}$	-0.074* (0.042)	-0.157*** (0.053)	-0.154*** (0.048)	$\ln y_{i,t-1}$	-3.006*** (0.727)	-2.849*** (0.920)	-2.534* (1.404)
Trade share		0.243*** (0.044)	0.237*** (0.043)	BR_{it-1}		0.053 (0.225)	-0.006 (0.288)
Government spending share			-0.175 (0.388)	Minority proportion %	0.070*** (0.015)	0.060* (0.021)	0.068* (0.039)
Investment share	0.117 (0.182)	0.295 (0.200)	0.317 (0.192)	Sec-school enrollment %	-0.074** (0.036)	-0.070** (0.029)	-0.086*** (0.031)
Sec-school enrollment	-0.190 (0.126)	-0.085 (0.114)	-0.055 (0.106)	Non-agri population %			0.007 (0.054)
constant	1.438*** (0.354)	1.843*** (0.390)	1.800*** (0.363)	constant	44.849*** (4.054)	42.76*** (10.815)	43.428*** (12.669)
Hansan J-stat	17.38	16.63	16.88	Hansan J-stat	7.94	14.39	18.18
(p value)	0.43	0.48	0.462	(p value)	0.34	0.496	0.253
Arellano-Bond test for 1 st order	-2.51	-2.55	-2.55	Arellano-Bond test for 1 st order	-1.92	-2.19	-1.97
(p value)	0.012	0.011	0.011	(p value)	0.055	0.028	0.048
Arellano-Bond test for 2 nd order	-0.12	-0.42	-0.55	Arellano-Bond test for 2 nd order	-0.36	-0.26	-0.14
(p value)	0.907	0.675	0.583	(p value)	0.723	0.797	0.892
Provinces	28	28	28	Provinces	28	28	28
Obs	136	135	135	Obs	136	136	136
# of instruments	26	27	28	# of instruments	16	25	26

Note: Heteroskedasticity consistent standard errors are reported in parentheses. *, **, *** indicate significance at the 10, 5 and 1% levels. The lagged variables refer to five-year lagged values (not averages). To control for the time effect time dummies are used for the grouped years 1978-1982, 1983-1987, 1988-1992, 1993-1997, 1998-2002. For brevity we dropped the time dummy estimates from the table. The birth rate is given on the basis of 1/1000. Secondary-school enrolment is the proportion of primary cohort entering secondary schools. Some provinces do not have the complete time series. All models are estimated using GMM-SYS method.

Table 3: Annual data models

Growth-birth model			Birth-growth model				
Dependent variable: $\Delta \ln y_{it}$			Dependent variable: BR_{it}				
	(1)	(2)	(3)		(4)	(5)	(6)
BR_{it}	0.007 (0.005)	0.014 (0.008)	0.004 (0.005)	$\Delta \ln y_{it}$	9.040 (5.337)	-4.168 (4.472)	-4.321 (4.409)
$\ln y_{i,t-1}$	-0.130*** (0.047)	-0.059 (0.035)	-0.183*** (0.033)	$\ln y_{i,t-1}$	0.732 (1.088)	-0.463 (1.150)	-0.499 (1.149)
D78	0.318*** (0.067)	0.020 (0.028)	0.021 (0.015)	D78	-9.263*** (2.714)	-6.325** (2.700)	-5.756** (2.750)
$D78 * BR_{it}$	-0.014*** (0.003)	-0.018** (0.007)		BR_{it-1}	0.260** (0.106)	0.308*** (0.090)	0.307*** (0.090)
$D6577 * BR_{it}$		-0.018*** (0.006)		$D78 * BR_{it-1}$	0.407*** (0.125)	0.287** (0.126)	0.288** (0.126)
$D65 * BR_{it}$			-0.016*** (0.005)	Famine* $\Delta \ln y_{it}$		14.100*** (3.711)	14.219*** (3.671)
Famine	-0.861*** (0.291)	-1.014*** (0.239)	-1.305*** (0.242)	$D78 * \Delta \ln y_{it}$			-5.939 (4.693)
$Famine * BR_{it}$	0.028** (0.011)	0.018 (0.012)	0.032*** (0.011)	Sec-school enrollment%	-0.070*** (0.022)	-0.041 (0.024)	-0.040 (0.024)
Trade share	0.153 (0.112)	0.075 (0.046)		Famine	2.572 (2.321)	0.799 (2.305)	0.851 (2.292)
$D78 * Trade\ share$			0.173* (0.097)	Constant	13.643* (6.679)	18.847*** (6.293)	18.960*** (6.249)
Government spending share	-0.496*** (0.176)	-0.357*** (0.093)	-0.454** (0.216)				
Investment share	0.094* (0.054)	0.189*** (0.064)	0.263*** (0.069)				
Sec-school enrollment	0.319*** (0.077)	0.101 (0.070)	0.339*** (0.071)				
Constant	0.520 (0.339)	0.426* (0.229)	1.183*** (0.240)				
Hansan J-stat (p value)	2.63 0.955	1.19 0.997	4.78 0.997	Hansan J-stat (p value)	7.37 0.391	9.94 0.192	7.28 0.401
Arellano-Bond test for 1 st order (p value)	-2.74 0.006	-3.22 0.001	-2.69 0.007	Arellano-Bond test for 1 st order (p value)	-3.27 0.001	-3.34 0.001	-3.33 0.001
Arellano-Bond test for 2 nd order (p value)	-0.26 0.797	-1.2 0.230	-0.17 0.864	Arellano-Bond test for 2 nd order (p value)	-0.77 0.438	-0.81 0.419	-0.80 0.426
Provinces	28	28	28	Provinces	28	28	28
Obs	810	810	810	Obs	951	951	951
# of instruments	34	35	42	# of instruments	30	31	32
Utilized lags	2-5	2-5	2-9	Utilized lags	2-5	2-5	2-5

Notes: See notes under Table 2. Lagged values are one-year lags. Time dummies are dropped for brevity.

Figure 1: Scatter plot growth of GDP per capita and Birth rate by province, 1954-2002

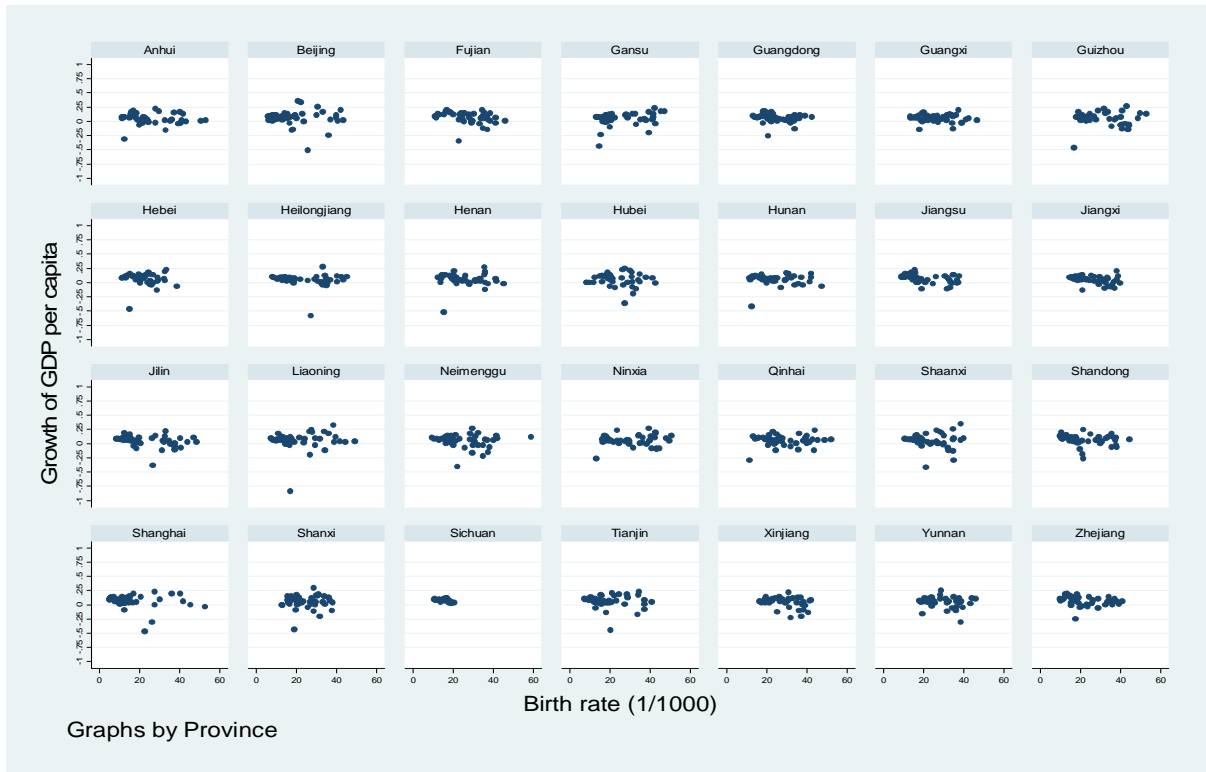


Figure 2: Scatter plot log of per capita GDP and birth rate by province, 1953-2002



Figure 3: Birth rate by province, 1953-2002 (vertical line indicating 1978)

