FERTILITY IN CHINA IN 2000: A COUNTY LEVEL ANALYSIS

A Thesis

by

HEATHER KATHLEEN MARY TERRELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2005

Major Subject: Sociology

FERTILITY IN CHINA IN 2000: A COUNTY LEVEL ANALYSIS

A Thesis

by

HEATHER KATHLEEN MARY TERRELL

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Dudley Poston (Chair of Committee)	Alex McIntosh (Member)
Don Albrecht (Member)	Rogelio Saenz (Head of Department)

May 2005

Major Subject: Sociology

ABSTRACT

Fertility in China in 2000: A County Level Analysis. (May 2005)

Heather Kathleen Mary Terrell, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Dudley L. Poston, Jr.

In order to maintain itself into the future, the People's Republic of China undertook in the 1970s a legendary demographic endeavor dealing with the artificial constraint of population growth. The "later, longer, fewer" policy and the more rigid one-child policy were efforts to expedite the demographic transition in the country. The ultimate goal was the stabilization and eventual decline of the population, via fertility at below-replacement levels for an extended period of time. According to the 2000 census, the total fertility rate (TFR) for China was 1.22—well below 2.1, the replacement level of fertility. However, the country's TFR fluctuated spatially with rates of .86, 1.08, and 1.43, for cities, towns, and rural areas, respectively.

Undoubtedly, China's family planning policy is largely responsible for the nation's current low fertility, as well as the geographical variation in fertility just mentioned. Research has shown, however, that other factors have played a part in this fertility transition and the subsequent variation at the regional, provincial, and county levels. In keeping with the expectations of demographic transition theory (DTT), quantitative studies conducted over the last twenty years have linked an assortment of socioeconomic factors with China's fertility decline and nationwide inconsistencies (Birdsall and Jamison 1983; Tien 1984; Poston and Gu 1987; Freedman et al. 1988; Peng 1989; Poston and Jia 1990; Poston 2000).

My thesis built on and extended the above work, using the newly available demographic data provided by Census 2000. I tested the efficiency of demographic transition variables in explaining the variation in the TFR among the counties of China by estimating twelve Ordinary Least Squares (OLS) regression equations. Specifically, I examined the ways in which variables such as ethnicity, agricultural detachment, urbanization, economic conditions, cultural norms and gender differences were related to Chinese fertility in a nationwide analysis and in two region-specific analyses. My results showed rather definitively that demographic transition theory is applicable for predicting and understanding fertility among the counties of China. Irrespective of the nation's extensive family planning policy, it is apparent that other factors contribute to the varying fertility rates across the country.

To my parents Kathy and Frank, for all their emotional and financial support,

to my companion Marshall, for his love and encouragement,

and to Lori who always kept up my spirits.

ACKNOWLEDGMENTS

First and foremost, I must thank my committee chair, Dr. Dudley L. Poston, Jr., who taught me everything I know about demography and China, and who encouraged and assisted me through my entire graduate school experience.

I am grateful to my committee members, Dr. Don Albrecht and Dr. Alex McIntosh, whose insights were of great help to me in writing this thesis.

For her contagious sense of humor when I needed it most, I have to thank my best friend Kellee. I would also like to express gratitude to my close friends in the department, D'Lane Compton, Chris Lewinski, Ron Lorenzo, Li Zhang, Amanda Baumle, Hua Luo, and Paul Min, for their support and assistance throughout this process. And finally, I would like to acknowledge our computer specialist in the department, Virgil Martinez, for all his help and patience.

TABLE OF CONTENTS

CHAPTER		Page
I	INTRODUCTION………………………………………………………...	1
II	LITERATURE REVIEW……………………………………………....	13
	Demographic Transition Theory………………………………….	13
	Chinese Fertility Studies………………………………………….	18
III	DATA AND METHODS…………………………………………………	34
	Data………………………………………………………………..	34
	Operationalization…………………………………………………	38
	Methods…………………………………………………………...	47
IV	DESCRIPTION…………………………………………………………	51
	Dependent Variables……………………………………………….	51
	Independent Variables……………………………………………...	53
	Transformed Variables……………………………………………..	62
V	REGRESSION ANALYSES OF THE TOTAL FERTILITY RATE…………………………………………………………	67
	Diagnostic Strategies………………………………………………	67
	Models 1-3…………………………………………………………	73
	Models 4-6: Regional Controls…………………………………….	77
VI	REGION-SPECIFIC REGRESSION ANALYSES OF THE TOTAL FERTILITY RATE……………………………………..	82
	Models 7-9………………………………………………………….	83
	Models 10-12……………………………………………………….	89
VII	CONCLUSIONS AND DISCUSSION…………………………………….	97
REFERENCES………………………………………………………………………		112
APPENDIX A………………………………………………………………………		122

Page

APPENDIX B……………………………………………………………………. 128

VITA…………………………………………………….................................. 134

CHAPTER I

INTRODUCTION

China, containing one-fifth of the world's population, has long recognized the need to slow down its rate of natural increase. In 1998, its rate of natural increase apparently dropped below 1%, the lowest rate in history excluding the 1959-61 period of famine (Attane 2001). Since China's borders are more or less sealed, its population growth is primarily determined by the total number of births and deaths. Thus, the successful deceleration of natural increase depends on the attenuation of the country's fertility rate. Although the Chinese government has played a large role in depressing the national fertility rate, other factors have also contributed to this transition.

The research to be reported in this thesis will analyze some of the socioeconomic and related factors that are associated with fertility among China's counties in 2000. Before discussing these issues, however, a short history of population and family planning in China will be presented. Hopefully these discussions will provide a perspective for the later discussion and presentation of this thesis.

While efforts to curb population growth were made in the 1950s and 1960s, the Chinese government took a more serious stance in 1971 with the delivery of the "wan (later), xi (longer), shao (fewer)" policy (Banister 1987; Bongaarts and Greenhalgh 1985; Cooney and Li 1994; Tien 1980). This strategy demanded that couples marry late, elongate their birth intervals, and have a small number of children. "Large families are undesirable" became the slogan of the day, and a family with two children was deemed

---

This thesis follows the style and format of *American Sociological Review*.

to be of perfect size (Tien 1980). Regulations were most restrictive in the cities where couples were urged to postpone marriage until age 25 for women and 28 for men and to have a maximum of two children (Attane 2002). Rural residents were given some leniency, with marriage age minimums of 23 and 25, and a family size limit of three children. A birth-spacing interval of no less than three or four years was required for both urban and rural couples. In 1977, a maximum of two children was directed at all couples, both urban and rural (Bongaarts and Greenhalgh 1985). "According to Chen Muhua, Vice Premier and Head of the Planned Reproduction Group of the State Council of the Central Government, China's overall rate of natural increase had fallen to 12.1 per 1,000 in 1978 from 23.4 in 1971, and the birthrate had dropped to 18.3 per 1,000 in 1978 from the previous high of 40.0 per 1,000 in the 1960s (Tien 1980: 65)."

In addition to the "wan, xi, shao" policy, important institutional changes were taking place in China at this time also. Following the demise of Mao Zedong in 1976, it became apparent that his administrative strategy was to blame for the country's deteriorating economy. As Prybyla (1990: 114) notes, "the signs of morbidity revealed by that economy were many, among them: chronic shortages of wanted goods; huge resource misallocations—waste; breakdown of the incentive system—low factor productivity; and technological backwardness." Peasants were disproportionately the targets of exploitation, paying high prices for manufactured items and being paid low prices for obligatory grain sales (Walder 1989; Lardy 1983; Riskin 1987). As a result, the disparity in income and food consumption among urban and rural areas grew larger (Walder 1989).

The economic reform effort began in 1979 and was entitled "Four Modernizations," indicating "the four areas—agriculture, industry, science and technology, and national defense—to be modernized (Prybyla 1990: 114)." Leaders were unwilling to surrender to the market system, so true economic freedom was never granted in this endeavor. Nevertheless, two decades of collectivized agriculture finally came to a halt with the initiation of the "household responsibility system," which converted commune farming into individual family farming (Feder et al. 1992; Prybyla 1990). Now commune land is leased to households for periods of up to 15 years, and the standard farm is three-quarters of an acre in size. Families still have to furnish the local state authorities with specific yields of certain crops at state-set prices, if they are to continue to be allowed to live on the land. After contractual quotas have been fulfilled with the government, peasants may consume any surplus food themselves or even sell it on the streets.

The Deng Xiaoping administration believed that substantial population growth would interfere with the goals of the "Four Modernizations," by "hampering attainment of full employment and by cutting into increases in capital accumulation, living standards, and education" (Chen 1979; Liu and Song 1981; Bongaarts and Greenhalgh 1985: 586). Projections regarding economic and sustenance resources suggested that China would boast nationwide security if it could shrink its population to 650-700 million by the middle of the twenty-first century (Liu and Song 1981; Bongaarts and Greenhalgh 1985). Their judgment hinging on these types of forecasts, the post-Mao leadership took extraordinary legal and administrative measures to control population growth.

Although the later-longer-fewer campaign was successful in reducing the fertility rate, by the late 1970s the concept of “shao” was modified to “one is enough” (Tien 1980). The issues which resulted in the one-child policy are not clear-cut. One large trigger was governmental awareness that the country’s young age composition would create growth for years to come if couples were allowed to have two children (Bongaarts and Greenhalgh 1985). By mid-1978 the decision had been made and in January 1979, the government instigated the renowned one-child policy—the most stringent national policy in the history of the world (Banister 1984; Bongaarts and Greenhalgh 1985).

Throughout 1979, the Chinese government became more resolute in their family planning endeavor, as they hoped to restrict the total population to 1.2 billion by the end of the 20th century. In 1980, Party Chairman Hua Guofeng announced that each couple of Han majority nationality would be restricted to one child only. A minute fraction of urban and rural couples (i.e. 5 and 10 percent, respectively) in exceptional situations would be permitted to have a second child (Bongaarts and Greenhalgh 1985). The policy focus was the solitary child; however, delayed marriage and delayed childbearing were still considered important factors and were to be given “appropriate stress” from local officials (Central Committee of the CPC 1983). In actuality these two issues have been accorded little emphasis in comparison with family size restriction (Bongaarts and Greenhalgh 1985).

The enforcement of this new policy was facilitated by earlier efforts, which made services and contraceptives widespread and easily attainable for the population at large. To promote acceptance of the one-child policy, the government established the "one child" certificate. This certificate guaranteed an array of benefits to couples with one child who pledged to have no more (Cooney and Li 1994). Three other factors strengthened conformity: (1) the 1980 marriage law making birth control mandatory, (2) the escalating use of abortion and sterilization as fall back measures to accomplish the population planning targets, and (3) fines against families having third and higher parity births (Banister 1987; Tien 1991; Bongaarts and Greenhalgh 1985). The estimated total fertility rate (TFR) fell from 2.7 in 1981 to 2.1 in 1984, creating enthusiasm among officials (Hardee-Cleaveland and Banister 1988).

Whereas the one-child policy has been very effective in urban areas, it has faced considerable obstacles in the country. In its initial stages, over 40% of rural births were unauthorized and thus in violation of the policy (Attane 2001). It's fair to say that a little bit of freedom went a long way for the effected peasantry. Economic reforms during the 1980s served to weaken the control of state and local powers over individuals, which stalled the performance of family-planning policies. The "household responsibility system" not only bestowed more autonomy on rural residents, it also increased incentives for large families by moving responsibility from the commune to the household. Additionally, state policies intended to strengthen the family planning campaign have often worked against it. One such case is the Marriage Law of 1980. "Although it increased the legal ages of marriage from 18 to 20 for women and from 20 to 22 for men, it invalidated the higher de facto age limits that had been in effect during the later-longer-fewer campaign of the 1970s, and effectively lowered the age of marriage" (Bongaarts and Greenhalgh 1985: 592).

To ease public criticism and defiance, the Party Central Committee released Central Document 7 in 1984, which called for more leniencies in regards to second births among rural couples and others with "real difficulties," provided that the 1.2 billion population goal for the year 2000 was respected (Bongaarts and Greenhalgh 1985: 588). They were optimistic that Document 7 would work "by open[ing] a small hole to close up a large one" (Greenhalgh 1986: 492; Qian 1997: 221). As head of the State Family Planning Commission, Wang Wei, remarked, cadres (i.e. family planning workers) "must not be too rigid in delivering a sermon to the masses" (Beijing Domestic Service 1985; Bongaarts and Greenhalgh 1985: 588). This sent mixed messages to cadres, who were expected to satisfy birth planning targets, and at the same time, accommodate the people (Hardee-Cleaveland and Banister 1988). Some cadres stood firm while others wavered, the result being geographical fertility fluctuations.

The watery plan proved ineffective as China's estimated TFR climbed from 2.1 in 1984 to 2.4 in 1986. Roughly one fifth of this upsurge was due to a change in age composition, while the rest was caused by a rise in fertility (Hardee-Cleaveland and Banister 1988).

Failed attempts at restructuring brought about another change of leadership in the State Family Planning Commission in the late 1980s, which then developed regulations to invigorate family planning in rural areas (Hardee-Cleaveland and Banister 1988; Qian

1997). Li Peng, interim Premier of the State Council, declared that "to keep the country's population at about 1.2 billion by the end of the century, China must 'strictly carry out' the policies of promoting later marriages and restricting each couple to only one child" (Hardee-Cleaveland and Banister 1988: 251). Provinces and localities were granted significant liberty in issuing second-child permits, though they were to remain stanch in their efforts to control population growth and abide by the central plan.

As Scharping (2003) has illustrated, these second-child permits vary extensively among the provinces. For instance, the most recent conditions for second births in urban Sichuan are: "(1) the first child is disabled or dead, (2) pregnancy after long years of childless marriage and a subsequent adoption, (3) in a remarriage one spouse has been childless, the other spouse already has had one child or two children, (4) one spouse or both Chinese spouses returned from overseas, Hong Kong or Taiwan, and (5) one spouse or both spouses are single children" (Scharping 2003: 100-101). Ningxia's conditions for second births are even broader: (1-5) same as above, "(6) one spouse or both spouses belong to a national minority with less than 10 million members, and (7) one spouse has been constantly working in underground mining for more than 5 years" (Scharping 2003: 100-101). This flexibility is needed since China is so large; however, it can also work against the fundamental objectives of the policy.

According to the 2000 census, the TFR for China was 1.22—well below 2.1—the replacement level of fertility. However, the country's TFR fluctuated spatially with rates of .86, 1.08, and 1.43, for city, town, and rural areas, respectively. The TFR varied from a low of .67 in Beijing to a high of 2.19 in Guizhou Province. While plunging birth rates in nearly all provinces suggest that program resistance is abating, widening intra-province differentials indicate that "pockets of resistance are hardening" (Attane 2001: 74).

Most investigations of China's TFR have been conducted at the provincial level, which cannot account for intra-province differentials. For instance, the 2000 TFR for urban Guizhou was 1.31, whereas the TFR for rural Guizhou was 2.42. When one takes into consideration all the one-child certificate holders, all the officially exempt second births, as well as all the unexcused out-of-quota births, it is easy to fathom how the TFR can vary across the country. The question is why? Why does the TFR vary? Do social, economic, political, cultural, and environmental factors play a part in this variation? Research indicates that the answer is yes—different dynamics influence whether family planning efforts are met with acceptance or resistance—yielding inconsistencies in the birth rate across China.

In order to understand the nation's current demographic conditions and to better evaluate the efficacy of China's family planning program, these disparities need to be located and analyzed. My thesis will explore this issue by examining the variation in the TFR among the counties of China. While family planning activities are vastly responsible for the fertility decline across China, my thesis will not include an analysis of the program itself. Rather, I will focus on socioeconomic indicators, which often lead to differences in policy enforcement and success (Poston and Gu 1987). I will generate and test a mixture of sociological and demographic explanations of the variation in the

TFR, drawing primarily from demographic transition theory (DTT). At present, this type of analysis has not been performed using the 2000 census data.

In past research, socioeconomic, political, cultural, and environmental factors were often associated with Chinese fertility, and some had more impact than others. Areas with the lowest fertility tended to be urban, with more economic development, and greater gender equality with regards to educational and occupational status. Remote border areas which rely on agricultural production, tended to have the highest fertility. These areas are rural, contain a large concentration of minority groups, and have greater gender disparity with regards to educational and occupational status.

I expect that my analysis will yield similar findings. My general hypothesis is that social and economic development is negatively related to fertility. Thus, I expect areas with greater levels of social and economic development to have lower levels of fertility. This hypothesis clearly echoes DTT, which posits that lowered mortality and lowered fertility arise out of socioeconomic development. As Minister Peng Peiyun stated herself in 1988, the population problem "is an important one, which is related to the overall economic and social development…" (Peiyun 1988; Hardee-Cleaveland and Banister 1988: 251).

I will now address specific socioeconomic variables and their hypothesized relationship with fertility among the counties of China. Variables such as ethnicity, agricultural detachment, urbanization, economic conditions, cultural norms and gender differences in illiteracy rates are expected to be related to the total fertility rate. Most of my expectations are in accordance with DTT or previous studies of Chinese fertility.

Cultural differences and frequent immunity from the one-child policy generate higher fertility among minority groups. Therefore, I hypothesize that counties with a lower percentage of minority population will have lower fertility rates.

The percentage of non-agricultural population in each county is expected to be negatively associated with fertility. In other words, a higher percentage of non-agricultural population should result in a lower total fertility rate. Other socioeconomic variables, such as urbanization and economic progress, are also expected to be negatively related to the total fertility rates.

Gender differences in illiteracy rates are anticipated to have positive associations with the total fertility rates. In other words, gender bias tends to result in higher fertility. Education is known to be negatively associated with fertility; thus overall illiteracy (males and females combined) is expected to be positively associated with fertility.

Census 2000 provides data on household structure (i.e. number of generations within a family unit) and divorce rates, which may expose the presence of traditional family norms. Traditional family norms and cultural values are expected to be positively related to fertility as well.

This thesis has several strengths. Data limitations caused problems and frustrations for Chinese demographers in past decades. I am fortunate to be working with Census 2000, which has been applauded for its addendum of short-form and long-form questionnaires. As mentioned above, there are two predictors which are intended to expose the presence of traditional family norms and cultural values. These predictors are the percentage divorced and the percentage of one generation family households.

Census 2000 is the first census to provide data on these issues; thus, previous studies of Chinese fertility have not been afforded such predictors in their models. I am eager to see their relative impact on fertility among the counties of China.

In anticipation of this thesis, I conducted a preliminary analysis of the variation in the TFR among the counties of Sichuan Province. Sichuan Province, commonly referred to as "the Texas of China," is geographically one of the largest provinces in the nation, encompassing an area of 485,000 square kilometers. This province contains 180 counties and has the largest population in China, with 87 million people. The analysis allowed me to identify and solve some methodological problems in advance, and also lent support to my hypotheses, serving as a spring board for my thesis.

Just as there are strengths, there are also weaknesses. My analysis could be vastly improved if it were possible to include an indicator addressing per capita income. Census 2000 does not ask such information, so I am limited in this way by the data. I would also like to incorporate a few family planning variables, since the family planning program has such a tremendous impact on fertility in China, apart from and in conjunction with socioeconomic factors. For instance, an independent variable pertaining to family planning costs might be especially helpful, since wealthier jurisdictions are at an advantage when it comes to birth restriction. Not only are their clinics better equipped, they are able to offer greater financial incentives to couples accepting the one-child certificate. Additionally, it would be interesting to include an indicator of abortion frequency, since abortions are often used to meet the birth planning targets. Census 2000 provides no data on family planning activities; so again, I am confined in this respect.

The subsequent chapter addresses pertinent literature regarding DTT, and reviews previous studies of Chinese fertility and their findings.

CHAPTER II

LITERATURE REVIEW

In this chapter, a detailed review of seven Chinese fertility studies, spanning nearly two decades of time, will be presented. Each of these studies investigated the impact of various socioeconomic factors on China's fertility decline and geographical fluctuations. Prior to reviewing these analyses, a brief account of demographic transition theory (DTT) will be provided, together with my rationale for using it as the theoretical framework in this thesis.

Demographic Transition Theory

During the late nineteenth and early twentieth centuries, there were a number of conjectures regarding the recent fertility declines in the West (Hirschman 1994). Some believed that the improved diets of industrial life had diminished the reproductive capacity of women (Notestein 1953). This theory fell short, since urban women who did not contracept could conceive as effortlessly as their ancestors two hundred years prior. Others attributed the decline to the creation of modern contraceptive methods. This interpretation also failed, since the movement toward fertility decline was already in progress before modern methods had any noticeable significance. Demographers, particularly Warren Thompson (1929), stressed the importance of social and economic factors in prompting fertility decline. These notions took shape over the following decades and materialized in the 1940s as Demographic Transition Theory (DTT)—chiefly the effort of Frank Notestein (Hirschman 1994; Notestein 1945, 1953).

In essence, demographic transition theory (DTT) posits that lowered mortality and lowered fertility arise out of socioeconomic development. As a society becomes industrialized, traditional belief systems are replaced by an alternative set of "motivating factors," which result in a succession of new behaviors, one being deliberate fertility control (Hirschman 1994; Notestein 1953).

According to Notestein, some of the motivating factors responsible for the fertility decline in Western Europe and North America were "reduced…pressures toward traditional behavior," "new opportunities for individual advancement," "education and a rational point of view," "the cost of child-rearing grew and…economic contributions by children declined," "falling death-rates…lowered the inducements to have many births," and that "women…found new independence from household obligations and new economic roles less compatible with child-rearing" (Notestein 1953: 16).

Various demographers have put different spins on DTT. For Kingsley Davis (1963), the key independent variable is the degree of household economic strain, which is determined by household size and impending economic resources. He believes that societies will engage in fertility control, along with a series of other behaviors, to relieve the economic strain associated with modernization and mortality reductions. Ansley Coale's (1973) perspective is more specific. He set down three conditions essential for the decline of marital fertility: "(1) a setting that allowed for fertility planning to be part of the calculus of conscious choice, (2) the availability of effective information about the means to control fertility, and (3) clear economic advantages of fertility control"

(Hirschman 1994: 212). These opinions illuminate an important point about DTT: "it has room for every causal variable" (Hirschman 1994: 212).

A number of empirical studies have endeavored to discredit DTT and for the last thirty years, it has faced considerable opposition (e.g. wealth-flows, economic, ideational, evolutionary, ecological). According to Knodel and van de Walle (1979: 219), "interpretations of the European fertility experience were based on a combination of hazy empirical impressions and theoretical preconceptions." Using a combination of micro, macro, and qualitative studies, these authors concluded that there was "no clear threshold of social and economic development required for the fertility transition to begin" (Knodel and van de Walle 1979: 225), since fertility declines appeared in countries with conflicting levels of socioeconomic development.

Caldwell (1976) also disapproves of our sustained reliance upon DTT, as it has undergone negligible renovations to accommodate the new body of knowledge in this area. His major criticism of DTT is the assumption that rationality accompanies industrial, urban society; hence, primitive societies are seen as irrational. As stated by Caldwell (1979: 116) "the movement from a society characterized by economically unrestricted fertility to a society characterized by economically restricted fertility is essentially the product of social, rather than economic, change." He posits that mass education lowers fertility, by producing a reversal in the flow of wealth. With education, female status goes up and children become more costly to raise; thus, the upward flow of wealth changes directions and begins to flow downward, from parent to child.

I agree with Caldwell's position that mass education reduces fertility by increasing female status and the cost of children. These educational affects are included in DTT and in other fertility theories that attribute low fertility to improvements in socioeconomic status. However, I do not necessarily believe that these affects induce a U-turn in the flow of wealth. Education is a given among variables negatively impacting fertility decline. The reasons it works the way it does can easily be argued, but the outcome is the same. I tend to agree with Hillard Kaplan's statement that "humans, like all other known organisms, invest in, rather than exploit, their offspring" (1994: 785); thus I tend to question Caldwell's wealth-flows theory of fertility decline. I believe that a shift occurs—children produce less than in the past and cost more—but historically, I do not think children have supported their parents to such an extent that there is a complete reversal (180 degrees) in the flow of wealth when the shift occurs. Additionally, in the London and Hadden (1989) study of Thailand, there was no attempt to measure the actual flow of wealth from child to parent or vice versa for obvious reasons—it is impossible to do. Empirically, the wealth-flows model suffers, just like any other theories of fertility decline.

My questions about the wealth-flows model raise two important points, which help to justify my reliance upon DTT for this thesis. (1) When researchers are investigating the utility of theories, their findings are guided by their hypotheses. Consequently, some empirical studies will find reason to question certain theories, such as DTT, because of their choice of predictors, models, or methods. (2) Fertility is very difficult to measure quantitatively and it is even harder to predict. Fertility theories which sanction more freedom in the way of predictors, therefore, stand a better chance of survival.

In reflecting on those points, DTT has endured a flood of criticism over the last three decades, but the theory "has ideas that are hard to ignore and …live on" (Mason 1997: 444). Classic DTT allows for a great degree of latitude because of the way in which it was originally crafted. The broad scope of DTT allows researchers to test out a variety of variables in a multitude of ways—a quality which has sustained its viability for half a century. "There is no consensus on an alternative theory to replace demographic transition theory" (Hirschman 1994: 214) because no other theories offer such elasticity. Thus, DTT remains the most widely accepted theoretical framework, and for the purposes of my thesis, will serve me best (Hirschman 1994).

As I noted earlier, DTT does not have a standard suite of predictor variables intended for fertility decline. Thus, the major challenge for researchers has been the selection of variables. This poses an even bigger problem for demographers studying China who are working with a very unique population undergoing an unprecedented fertility transition. The fertility reductions that have occurred in industrialized countries may well be fundamentally different from the reduction that has occurred in China. One obvious difference is that China is still a developing country, with much of its traditional belief system intact, yet it displays a TFR of 1.22. The second distinction is the one-child policy. While the policy is not solely responsible for the fertility reduction, it has been important in why the TFR is so low. Since the 1980s, researchers have worked to identify the variables most significantly associated with China's fertility decline and geographical fluctuations. I will now provide a detailed overview of these studies and their findings.

Chinese Fertility Studies

Birdsall and Jamison (1983) were the first demographers to address China's fertility at the provincial, prefectural, and county levels. In performing this study, they hoped to shed light on whether or not future economic growth and freedom could be expected to strengthen or weaken the government policy to reduce fertility. They suspected the effects of socioeconomic conditions and government programs to interact with one another; thus, they believed family planning programs would be more successful in places where socioeconomic conditions bolstered government objectives.

Due to a lack of data, they could not test the extent to which birth planning inputs triggered fertility change in different regions of the country. However, they did have data on which of the 29 provinces, autonomous regions, and municipalities had activated an explicit program of incentives and disincentives to deter births by 1979. Using cross-sectional data on income and other markers of socioeconomic development, and on incentives, they estimated a series of Ordinary Least Squares multiple regression models. To the degree that the birth planning inputs were uncorrelated with the socioeconomic markers and with the incentive variable, their effects were integrated into the unexplained residual.

Their multivariate study of the crude birth rate engendered three main conclusions which I will now address. First, they found that different levels of fertility are associated with different levels of development. China adheres to the common pattern: high-income areas are associated with lower levels of fertility.

The rural income variable explained 45 percent of the variance in crude birth rates within a single province and income explained between 17 and 38 percent of the variance in the crude birth rates across provinces. According to the authors, it is not income as such that reduces fertility. Rather, there are probably a number of unspecified variables that are correlated with income—like educational status and wage rates and career prospects for women—that more immediately impact fertility by raising the cost of children to parents. They believe that if they could control for these variables, the effect of income on fertility would be positive. These results indicate that levels of development have considerable influence on China’s fertility, apart from the impact of the one-child policy. Additionally, once income was controlled, urban variables had only a weak negative impact on fertility. This seems to suggest that it is not urban residence per se that lowers fertility, but rather the higher income associated with urban residence. Conversely, there may be more under-reporting of births in rural areas, which conceal the true rural-urban differences in fertility. The incentive variable showed no direct effect on fertility, most likely because the program was only a few years old at the time of this study.

Their second conclusion was that higher income not only resulted in lower fertility levels, it also engendered a greater degree of fertility decline between 1975 and 1979 among the counties of Jiangsu Province.

For their third and final conclusion, they discovered that the negative association between income and fertility weakened between 1975 and 1979. The authors put forth a couple of explanations for this finding. For one, crude birth rates were being depressed into a smaller range as they approached replacement-level fertility, thus reducing the explanatory strength of income. Secondly, income may have played a larger role in reducing fertility prior to 1975 than it did between 1975 and 1979, since policy enforcement became more influential in the latter period. Serious government efforts did not begin until 1972, and the initial effects (1972-75) of those efforts may have been more substantial in high-income areas where conditions were more favorable for fertility decline.

Tien (1984) carried out a succession of provincial level cross-sectional studies for the years 1978, 1979, 1980, and 1981, which investigated the relative impact of population planning programs and socio-economic development on China's fertility transition. He put forth four hypotheses to be tested: "(1) The higher the level of urbanization in a province, the lower is its fertility; (2) The higher total output per head in a province, the lower is its fertility; (3) The higher the grain output per head (i.e. higher economic well-being in an agrarian society) in a province, the lower is its fertility; (4) The higher the life expectancy at birth in a province, the lower is its fertility" (Tien 1984: 393). According to Tien, grain output per head is undoubtedly the most revealing marker of individual and family welfare in China, since the majority of the population is reliant upon agriculture for income and survival.

His provincial level analyses were disadvantaged by data problems and deficiencies. Still, Tien was able to make a few generalizing statements about China's fertility. Fertility varied extensively but stayed high in the provinces prior to 1970, before the population planning measures were amplified. Fertility began to decline after 1970, but the progress was uneven during the period of program intensification. The fertility differences among provinces prior to 1970 grew larger after the program augmentation, especially amid the urban and rural provinces. Fertility appears to have increased between 1979 and 1981, probably because of transformations in age and marriage structure.

I will now address Tien's results pertaining to socioeconomic variables and differential natural increase. Prior to 1981, grain output per head had little influence over natural increase. However, grain output per head was negatively and significantly associated with natural increase in 1981, after the changes were made which returned productive control to peasant households. Tien speculates that where grain output is deficient, less money can be allocated to agricultural mechanization and old age retirement; thus, the motivation to have more children is greater in those places.

He also found that urbanization, output per head, and life expectancy were all negatively and significantly related to natural increase. Tien believes the correlations between the socioeconomic variables and natural increase would have been even higher, if the life expectancy variable had been eliminated from the model. Given that monetary benefits are often used to reward couples who abide by the policy, Tien reasons that wealthier provinces have an advantage over the others with regards to fertility restriction.

Consequently, most of these findings lent support for his hypotheses. Both socioeconomic change and government programs were crucial to China's fertility transition, and justifications of this shift merit the inclusion of both factors. Tien's findings challenge the notion that China's fertility decline was entirely the outcome of an extensive family planning program.

Poston and Gu (1987) explored the interactions between socioeconomic development, family planning, and fertility among China's 28 subregions, around 1982. In their model, they included 17 development variables, which exposed different social and economic characteristics of the subregions. They grouped these variables into four separate indexes relating to structural development, female status, quality of life, and rural quality of life. They also incorporated 11 family planning variables; ten were in regards to participant behavior and one was in regards to expenditures. Their dependent variable was the total fertility rate.

Their main objective was to determine the extent to which socioeconomic factors and family planning behavior influenced fertility change and decline among the subregions. They hypothesized that both socioeconomic development variables and family planning variables would be negatively associated with fertility among the subregions during this period of analysis. They also hypothesized that socioeconomic development variables would have an indirect negative effect on fertility via the family planning variables. Their results substantiated these conjectures. All of the variables were negatively associated with fertility, although their relative effects differed appreciably.

Poston and Gu found that certain aspects of socioeconomic development and family planning behavior had large direct negative effects on fertility. Additionally, they discovered that some socioeconomic factors had indirect effects on fertility via family planning behavior, the largest of which was the female status index. In regards to the family planning variables, they found that contraceptive behavior had a large direct negative effect on fertility, while family planning costs had a relatively weak effect on fertility.

Poston and Gu's findings support earlier studies (Tien 1984; Birdsall and Jamison 1983) suggesting that socioeconomic development has a direct effect on fertility, apart from family planning efforts. Their results also correspond to other studies (Kelly et al. 1983; Mauldin et al. 1978) showing that specific aspects of development effect fertility through contraceptive behavior.

In brief, their results indicate that China's fertility transition should not be attributed exclusively to family planning programs. Socioeconomic development has had a hand in this shift, affecting fertility both directly and indirectly, and these effects should be considered in any empirical investigation of China's fertility decline.

While the Chinese government has been successful in circulating fertility policies throughout the country, its policy control is indirect and must operate through local units (Freedman et al. 1988). Thus, the initiation, enforcement, and success of these policies are highly dependent upon local units of control. Due to wavering local conditions and resource availability, the ultimate outcomes of central policies are greatly inconsistent, not only across, but within provinces. Freedman et al. (1988) wished to throw light on this issue by examining reproductive behavior in local communities, simultaneously revealing the degree of variability in fertility within provinces. They performed a vast investigation of China's fertility 1973-1982, probing data from 815 production brigades (now administrative villages).

Using data from the 1/1000 *Survey* for China, they constructed and analyzed three-year fertility rates for each woman in the sample. They studied four provinces in detail (Hebei, Henan, Liaoning and Sichuan), which differ noticeably in socioeconomic development, and which constitute one-fourth of China's population. Efforts were centered on rural, rather than urban units, since rural units are more plentiful and vary more by education and fertility levels.

They discovered that markers of reproductive behavior fluctuated appreciably among different production brigades, especially in rural areas. For instance, there was substantial deviation among the rural production brigades in age at marriage. In Hebei alone, the proportions married prior to age 23 contrasted from around 15 percent to 95 percent. Also relevant, were the proportions of first births taking place after the initiation of the one-child policy. "The percentage of production brigades in which the proportion of first births was 40 or more is 75 for Hebei, 46 for Henan, 64 for Liaoning, and 73 for Sichuan" (Freedman et al. 1988: 44).

Additionally, they examined the predominance of various contraceptive methods in the four provinces and in the country as a whole and found greater variation between production brigades than between provinces. Regarding abortion ratios (100*abortions/births), there was significant variation between and within provinces, and the highest abortion ratios occurred in the urban sectors of each province.

The single explanatory variable accessible from the 1/1000 Survey was education of women. They found evidence of macro and micro educational effects on fertility, although these effects varied among provinces and transpired at different times. It seems that in Liaoning, the educational milieu of the units had an enormous impact on the women's fertility level, irrespective of the women's own education. The authors believe these macro educational effects may have been a consequence of other variables not included in the study or an outcome of using expansive educational categories. The urban fertility rates of Liaoning and Sichuan were much lower than those of the rural units at every educational level, which suggests the family planning program was already operative in urban areas between 1973 and 1976.

In 1973-6, the usual negative relationship between individual education and total fertility was not observed with any uniformity for local areas, and it was not observed at all in urban areas. Even more salient is the finding that the macro-educational effect was gone by 1979-82. In other words, the fertility of rural women was essentially unaffected by the illiteracy level of their brigade. Moreover, the negative association between individual education and fertility existed in just two out of the four rural micro-education groups, and did not exist at all in urban areas.

Their study suggests that China's family planning program has surmounted many obstacles caused by low educational levels; however, micro-education disparities in fertility have continued to exist in rural areas. There is extreme variation among the production brigades in terms of illiteracy levels—ranging from just 3 percent illiteracy to between 80 and 100 percent. This variability is important to consider, since educational levels are often related to other socio-economic variables. Most noteworthy, is that this study did not champion the notion that China's demographic changes are entirely the outcome of a family planning program.

China's TFR fell by over 50 percent in one decade (1970-1980); however, the timing and pace of the transition was erratic across the country. Peng (1989) believed a great deal of this variation could be explained by way of the major determinants (e.g. socioeconomic and institutional aspects) of the country's fertility transition. Accordingly, he used the 1982 national fertility survey to examine specific socioeconomic effects on fertility among China's 28 provinces.

Peng tested the effect of education on fertility and found that the number of children ever born to women drops continuously with rising education. Women with zero education always had the greatest number of children, whereas women with high school education or better had the smallest number of children. The mean parities of illiterate women were 20 percent higher than those with a junior middle-school education and 36 percent higher than those with a high school or higher education.

As one would expect, the number of children born to urban women was lower than the number born to rural women. Even when education and province location were held constant, Peng found urban women aged 35 to have 30 percent less children than their rural counterparts. Moreover, the effect of rural-urban on fertility appears stronger for the younger age groups than the effect of education.

Peng also found occupation to be a key determinant of women's reproductive behavior. He compared fertility among four occupational groups (i.e. peasant, worker, cadre, and housewife) and uncovered large occupational discrepancies. Working women and cadres had fewer children than peasant women and housewives, which could be attributed to differences in work schedules. Generally, peasants had more children than any other group, whereas cadres had fewer children than any other group.

Peng discovered that fertility levels vacillated among provinces, primarily between the eastern and western parts of the country. Fertility tended to be lowest in the east, highest in the west, and approximately the national average in the middle section of the country.

For the most part, these findings are analogous to previous studies. The family planning program "has not operated in a vacuum;" many socioeconomic factors have influenced and continue to influence China's fertility shift.

Poston and Jia (1990) were among the first demographers to conduct an extensive county level Chinese fertility analysis. Using data from the 1982 Census of China they investigated and compared the impact of four independent variables (i.e. economic development, the infant mortality rate (IMR), the illiteracy rate, and the percentage of industrial employees) on the general fertility rate (GFR) in 2306 counties. Their goal was to determine the influence of pertinent socioeconomic factors on fertility transition, and the degree to which this influence varied among counties in different regions of the country.

Their dependent variable was the general fertility rate (GFR). In evaluating its means and standard deviations, they discovered that there was considerable variability about the average GFR value of 98.5. It ranged from a low of 36.2 in Shihizhi city in the Xinjiang autonomous region to a high of 232.5 in Butuo county in Sichuan province. They also found that the GFR in 1981 varied much more among the counties of China than among the provinces. As expected, the average GFR in rural counties (102) was substantially higher than in the urban counties (68).

As mentioned above, four independent variables were used. The first is an economic index, which signifies the "gross value of industrial and agricultural output per head (in yuan) in 1982" (Poston and Jia 1990: 508). The range of this index is rather impressive—extending from a low of 70Y in Haiyuan county in the Ningxia autonomous region to a high of 28,475Y in the Yanshan district of Beijing.

The second predictor is the infant mortality rate (IMR) in 1981, which had a mean value of 39.0 (SD 29.1) among the 2,306 counties included in the study. The IMR ranged from a low of 6 in Langfang city in Hebei Province to a high of 319 in DaQaidam county in Qinghai province.

The third predictor is the illiteracy rate, "the percentage of the population aged 12 or over in the county in 1982 that is illiterate or semi-literate; a literate person knows at least 1500 Chinese characters, can read simple books and newspapers, and can write simple messages" (Poston and Jia 1990: 510; Beijing Review 1984). This independent variable had an average value of 34.1 (SD 13.9), ranging from a low of 2.7% in Lianjiang county in Guangdong province to a high of 84.9% in Dongxiang county in Gansu province.

The fourth and final independent variable is the percentage of industrial employees in the labor force in 1982. This predictor had a mean value of 13.3 (SD 14.6), spanning from a low of 0.1% in Jishishan county of Gansu province to a high of 76.9% in the Yanshan district of Beijing. All four of the independent variables showed great variation among the counties of China; subsequently, they influenced fertility at varying degrees. I will now address their findings.

Of the four socioeconomic markers, they found the IMR to have the largest and most significant effect on fertility amid all of the counties, and amid the urban and rural counties evaluated independently. Their investigation of the rural counties showed that the four independent variables have different effects on different regions of the country. In Dongbei and Xibei, the IMR has the most significant impact on fertility, in Huabei and Xinan, the illiteracy rate has the most significant impact, and lastly, for Huadong and Zhongnan, the economic index has the most significant impact on fertility. These findings reinforce prior provincial level analyses linking socioeconomic factors to Chinese fertility trends.

Poston (2000) conducted cross-sectional studies of China and Taiwan using subregional data at two time periods, the early 1980s and 1990. His main purpose was to investigate the degree to which social and economic development stimulated fertility reductions in each of the countries. Both China and Taiwan have undergone drastic fertility transitions over the past 40 years—their TFRs dropping from 6 in the early 1950s to less than 2 by 1997. Yet, different sets of events may have triggered the transitions for each of these countries. Taiwan, as a whole, experienced the transition earlier than China. Many researchers believe that Taiwan's fertility decline was due largely to socioeconomic development, since involvement in the country's family planning programs has always been voluntary. Conversely, China's fertility decline is believed to have been generated by both socioeconomic development and its family planning programs, since involvement in the latter has been more or less mandatory.

He explored the link between socioeconomic development and fertility "by using data for the counties and provinces of China for 1982, 1990, and 1995; and for the counties and cities of Taiwan for 1980, 1990, and 1995" (Poston 2000: 41). This study is different from other investigations of this nature because of its subregional focus.

In conducting his analyses, Poston regards four factors as having especially significant effects on fertility: "(1) advances in economic development, specifically increases in economic productivity and participation in the nonagricultural labor force; (2) improvements in general health conditions, especially reductions in infant mortality; (3) improvements in social conditions and livelihood, particularly in educational attainment; and (4) absolute and relative improvements in female status" (Poston 2000: 43). Previous studies have found links between one or more of these factors and fertility decline in China and in Taiwan. I will now provide a synopsis of his analyses and findings for China.

In his analysis of the fertility trends among China's counties in 1982, the general fertility rate (GFR) served as the dependent variable. The GFR had a mean value of 98 in 1981 among the 2,300 counties included in the study, ranging from a low of 36 to a high of 233. He found three out of his four independent variables—the infant mortality rate (IMR) in 1981 (+), the illiteracy rate in 1982 (+), and the percentage of industrial employees in the 1982 labor force (-)—to be significantly related to fertility in the hypothesized direction (in parentheses above). The infant mortality rate was the most influential predictor and his model accounted for nearly 43 percent of the variation in the GFR in 1982.

The general fertility rate (GFR) again served as the dependent variable in Poston's 1990 analysis, and as in 1982, varied greatly among the counties of China. The GFR had an average value of 91, spanning from a low of 35 in Jiading District in Shanghai to a high of 210 in Cuoqin County in Tibet. The results for the 1990 multiple regression equation were similar to those for 1982. Illiteracy and the crude death rate (CDR) had significant positive effects on fertility, while industry had a significant negative effect. Illiteracy and industry were the most important predictors and his model accounted for 38 percent of the variation in the GFR in 1990.

The results from his provincial level analyses also indicate that social and economic development has played a role in determining Chinese fertility patterns. Poston hypothesized that the illiteracy and infant mortality variables would be positively related to fertility, whereas the income and life expectancy variables would be negatively related to fertility. All four of his equations showed income and infant mortality to be significantly related to fertility in the hypothesized direction. The effect of income was strongest in two of the equations, while the effect of infant mortality was strongest in the other two. The life expectancy and illiteracy variables did not excel in any of the models, an issue possibly attributable to measurement problems.

Although Poston faced various multicollinearity problems regarding the predictors in his fertility analysis of Taiwan, he found evidence that social and economic development were important factors in the country's fertility decline.

According to these results, socioeconomic factors have impacted fertility in China and Taiwan to comparable degrees, in spite of the fact that Taiwan is more developed than China. Poston credits China's extensive fertility-control polices, however, for bringing its fertility rate down to the level of Taiwan's over such a short period of time. His findings underpin prior analyses relating socioeconomic factors to Chinese fertility patterns, and also lend support to demographic transition theory of fertility decline.

The aforementioned studies, spanning nearly twenty years time, have established various connections between aspects of socioeconomic development and fertility decline in China. Although scholars often question the soundness of DTT, I agree with Poston (2000: 58) that "these results provide confirming evidence that a social and economic development-based theory of fertility decline, as represented by the broader theory of demographic transition, continues to have considerable relevance for understanding variation in fertility among the subregions of China."

So as to determine which aspects of socioeconomic development wield the most power over fertility in China, it is essential to analyze fertility trends across the nation. This thesis is an analysis of the variation in the TFR among 2,873 Chinese counties, based on the 2000 Population Census of China. In accordance with previous studies of Chinese fertility, I will consider the contribution of both socioeconomic status and gender differences in shaping fertility trends. I will also examine the relationship between traditional family norms and fertility. In the next chapter, I will present a discussion of the data and methods to be used in this thesis.

CHAPTER III

DATA AND METHODS

Chapter III provides a thorough description of the 2000 census of China, and also addresses the benefits and downfalls of using it as the solitary data source for this thesis. Additionally, the dependent variable and each of my independent variables will be operationalized. In doing so, I explain my reasoning for including them in the study and also attend to individual measurement issues. And finally the statistical methods to be used in my fertility analysis will be dealt with one at a time in considerable detail.

Data

The data source for my thesis is the 2000 Population Census of China—"arguably the world's most ambitious census ever" (Kennedy 2001: 1)—entailing "10,000 tons of paper for questionnaires, 5 million enumerators, and a million supervisors" (Lavely 2001: 755)."

Work on the census began at 00:00 hours on November 1st in the year 2000, which by the Confucian calendar was the year 4698. This was the Fifth National Census of the People's Republic of China and the first count in ten years. Government officials, teachers, and volunteers were utilized in the endeavor, visiting over 350 million families across the country (Landsberger n.d.). Ninety percent of the families were interviewed using a short-form questionnaire and each exchange lasted approximately ten minutes (Weaver 2000). Encounters were more time-consuming for the other ten percent who were interviewed using a long-form questionnaire. While the census period was scheduled to last ten days, 10 million missing Hunan residents compelled the Chinese Government to prolong the project by almost a week (Daily Reports 2001).

One of the most constructive changes in Census 2000 is that the data cannot be used to evaluate the functioning of local governments (Chan 2003). Generally, government statistics are used to assess policy implementation, which creates a situation that is inhospitable to accurate reporting. The State Council decision to safe-guard the data in this way was an effort to secure a more valid account of China's current demography.

Census 2000 has been commended for its use of short-form and long-form questionnaires, which provide rich data for a wide-range of prospective research projects (Chan 2003). Moreover, it is the first census in thirty years to introduce a "zanzhu renkou ('temporary population') form to record this population (the length of stay was reduced to less than six months), whereas the remaining, 'ordinary resident' population (changzhu renkou) was recorded in the regular census form" (Chan 2003). This incorporation will be beneficial for my thesis research because it has been hypothesized that many out-of-quota births are children of the "floating population."

Census 2000 has also been highly praised for using a more liberal definition of "urban" than was used in the past (Lavely 2001; Kennedy 2001). Whereas earlier censuses automatically counted village committees as rural, Census 2000 avoided such robotic classifications. The urban definition employed in the 2000 census bestows village committees urban status on the basis of density measures and other benchmarks (Lavely 2001). Furthermore, rural migrants who had resided in an urban area for over six months were labeled urban residents in the census, regardless of their household registration status (Kennedy 2001). The 2000 census reveals a national urbanization rate of 36 percent—five percentage points higher than the rate obtained from urban registration data in 1999 (31 percent). This means Census 2000 captured 67 million more urban residents on account of the broader classification system (Lavely 2001; Kennedy 2001).

One downfall of Census 2000 is its estimated undercount rate of 1.81% (22.46 million), a rate higher than that of the previous censuses (Chan 2003). Experts have attributed this problem to a variety of factors. First, it could indicate that the population is more "diverse and mobile" than it was in previous censuses. Additionally, there may be more under-reporting of "above-quota birth" children, as well as homeless and nomadic sex workers. Third, it could be that many people gave erroneous answers or omitted some of the questions, either advertently or inadvertently (Chan 2003).

The superior quality of former Chinese censuses stemmed from a rigid household registration system and strong ties connecting the registers with the census (Lavely 2001). The connection between registration and census work is overt and direct. During the period immediately preceding a census, the household registers are revised and authenticated in a process called "rectification" (*zhengdun*) (Lavely 2001: 764). The local officials then hand the polished registers over to the enumerators, who use them to put together a list of households in the enumeration districts. Since the enumerator is provided with information on each household prior to the household interview, the interview is a rough check of the registration information.

Census 2000 was performed under less auspicious circumstances. The household registers are not as accurate as in the past because policy requirements have led to an extensive falsification of local population conditions (Lavely 2001). For instance, pre-census rectification uncovered 50,000 and 130,000 deceased persons left on the registers in Wenzhou city and Chongquing Municipality, respectively. While these mistakes were rectified, “it is far easier to purge registers of the dead than it is to add the living” (Lavely 2001: 764). Above-quota birth children are frequently excluded or hidden from the register since both citizens and officials are motivated to conceal them. Parents hide surplus births to evade fines, while local birth planning officials (cadres) conceal them to convey the appearance of successful job performance (Lavely 2001).

In order to safeguard census work “census officials directed that all unregistered out-of-plan children be registered, that no fines be collected from parents of these children, and that officials be granted amnesty for previous birth planning falsifications as long as they were truthfully reported in the census” (Lavely 2001: 764). Similar guarantees were made before the 1990 census, only to be broken soon after (Kennedy 2001). Chinese remembered these infringements, which prevented them from coming forth in 2000. What is more, local reports suggest that many who did disclose excess births in 2000 were indeed penalized. “In some cases, the fines were reduced or eliminated, only to be replaced by ‘child registration fees’,” as indicated by a follow-up in Beijing (Kennedy 2001: 2).

Undoubtedly, there was a great propensity for error in Census 2000. Rumors of gigantic undercounts in the media, however, were probably overstated and due to confusion surrounding the enumeration of out-of-province migrants (Lavely 2001). “The enormous amounts of data from the Census 2000 are no doubt a gold mine for interested scholars, but there will also be more work, and more challenges, ahead if we wish to use them properly and fully” (Chan 2003: 7).

Results from the 2000 census indicate that China maintained its position as the most populated nation with 1.266 billion persons, or roughly 21 percent of the earth’s population (Lavely 2001; Kennedy 2001). The government’s target ceiling of 1.3 billion for the year 2000 was realized.

The country’s population has matured slightly over the last ten years—the proportion 65 and older increasing by 1.4 percent (5.6 to 7.0). Nevertheless, China remains quite young and boasts a sizeable workforce, as 70 percent of the population falls into the 15-64-year age bracket (Lavely 2001).

Operationalization

This thesis will analyze one dependent variable among the 2,873 Chinese counties: the total fertility rate (TFR). The TFR is calculated by adding up the age specific fertility rates (ASFRs), and multiplying the sum by the width of the age interval of the ASFRs. Typically, ASFRs are based on five-year intervals, so the sum is multiplied by 5. The TFR value can be understood as signifying the total number of live births an artificial cohort of 1,000 women would end up having if, as they survived through their 35 years of childbearing, they were exposed to the ASFRs of a county in China in 2000.

Previous Chinese studies have utilized a variety of fertility measures, including the crude birth rate (CBR), the general fertility rate (GFR), and the TFR. The GFR and TFR are usually preferred over the CBR since they are more statistically refined. The CBR stays true to its name, given that its denominator is the entire midyear population, which contains a large number of people who are not at the risk of childbearing. The GFR is superior to the CBR in this respect because its denominator is limited to women of childbearing age (the numerators of these two rates are identical: the number of births in year *z*). Still, the GFR suffers because it does not account for age-related fertility fluctuations. That is to say, women in their 20s will usually have higher fertility rates than women in their 40s. If at all possible, we should allow for these disparities in our fertility calculations. The TFR satisfies this need because it is based on the ASFRs, and when data is permitting, demographers typically rely on the measure for this reason.

In recent years, there has been some skepticism regarding the effectiveness of the conventional TFR in lowest-low fertility populations. According to Sobotka (2004) and others (see Kohler, Billari, and Ortega 2002), the postponement of fertility is depressing the period fertility rates in the lowest-low fertility countries. This tempo effect distorts the conventional TFR in that there is a supposed "recovery" of the TFR at the end of the fertility postponement. Thus, Sobotka believes lowest-low fertility is caused by increasing age at motherhood—a temporary trend that will disappear once postponement comes to a halt. I concur that this is problem in many European populations where the mean age at first birth is very high (i.e. 28-30 years of age); however, I do not believe it poses much of a problem in China since childbirth generally occurs shortly after marriage (i.e. 23-25 years of age). I am confident that the TFR is the superior choice of measurement for my county-level analysis of fertility in China.

Earlier analyses have incorporated an assortment of independent variables to capture the influence of socioeconomic development and modernization on fertility patterns in China (Birdsall and Jamison 1983; Freedman et al. 1988; Peng 1989; Poston 2000; Poston and Gu 1987; Poston and Jia 1990; Tien 1984). It is important to remember that, while variable selection is guided by theoretical concepts, it is constrained by the information provided in the data source. Data limitations created many obstacles for demographers studying Chinese fertility in previous decades. Census 2000 was developed to address several pressing social and economic concerns in the country; thusly, I have more freedom in the way of variable selection. Nevertheless, the confines of the data prevent me from including several pertinent variables in this thesis.

I will now address the operationalization of the ten independent variables in my study. All of my independent variables concern the 2,873 counties of China in the year 2000. [For sake of simplicity, I will not restate this for each independent variable.]

Three variables measure urbanization and economic progress among the counties of China: the percentage of the total population that is urban (i.e. residing in an urban area as defined by the census), the percentage of the total population that is employed in non-agricultural related work, and the percentage of the total employed population that is employed in white-collar work (i.e. professionals, associate professionals, managers, and administrators). As maintained by DTT and a myriad of demographic studies, urbanization is positively related to socioeconomic status, and both urbanization and socioeconomic status are negatively related to fertility. These independent variables attend to these issues nicely, by measuring the variation among the counties in levels of development and modes of employment. I expect all three of them to be negatively related to fertility.

Three variables measure the presence or absence of rudimentary education in an area: the percentage of the total male population aged 15 and over that is illiterate, the percentage of the total female population aged 15 and over that is illiterate, and the percentage of the total (male and female) population aged 15 and over that is illiterate. Since education has been shown to have a negative impact on fertility, I anticipate that each of these illiteracy variables will be positively associated with fertility.

I employed the male and female illiteracy variables to create a female status variable measuring educational inequality. The percentage female illiterate divided by the percentage male illiterate will produce such an indicator, since it will expose gender bias in regards to basic educational standing. Because women's status has been found to have a negative relationship with fertility and since inequality lowers the status of women, I expect that my female status variable will be positively associated with fertility.

It is common knowledge that rural areas have higher fertility rates than urban areas in China. How do the seven variables I just mentioned play a part in the variation of fertility among urban and rural areas? These variables not only influence fertility, they influence each other, which give rise to different dynamics, which change living conditions, which ultimately influence fertility again. In other words, they not only have a direct effect on fertility, they have an indirect effect on fertility via each other and outside forces. I will now provide a little background in an effort to elucidate the urban-rural fertility gap.

China's rural population numbers 866 million people. The current one percent annual growth rate means we can expect it to increase by about 9 million per year (without rural-to-urban migration) (Johnson 2000). Since rural areas have had higher birth rates in the past and are still slightly higher today, the rural labor force is overwhelming the job supply in the countryside. Even as rural fertility rates continue to drop, this situation will worsen over the next few years on account of "population momentum".

Urban residents, especially in China, tend to have elevated education and income levels in comparison to their rural counterparts. This is because there are simply more opportunities for individual advancement in urban areas. According to Johnson (2000: 332), "the major policy change that can be made to facilitate the required adjustment in the rural labor force is to equalize the educational opportunities between rural and urban areas." As it stands now, government officials are interfering with their own family planning and economic objectives by dumping a poor educational policy on the rural population. Rural migrants searching for work in urban areas are disadvantaged because they lack the education necessary to attain employment.

China's cities cannot accommodate an ever-expanding, unemployed rural population; therefore, another favorable policy decision would be to invest in the development of rural areas (Johnson 2000). Economic reforms over the last twenty years have generated more nonagricultural jobs in rural areas, but it is insufficient when you consider that there are millions and millions of people who will soon be flooding the labor market. This situation may well exacerbate an already substandard rural economy.

In the case of China, the inferior socioeconomic conditions of rural areas help to explain their higher rates of fertility, but there is one more piece of the puzzle: son preference. Rural couples are more likely to violate the one-child policy in an effort to achieve a son, whereas urban couples are not. Son preference abounds throughout China, so why this contradiction? While there may be social pressure, there is less financial pressure for urban residents to produce sons. Urbanization has increased the educational and occupational standing of women, and in turn, the status of women has risen. Furthermore the government is more generous with urban residents in terms of education and other key subsidies. Quite the opposite, rural residents are primarily dependent upon sons for old age security.

All of the independent variables I've cited influence fertility both directly and indirectly, promoting an urban-rural fertility gap. I hope I was able to illuminate some of these interlocking issues for the reader.

As I mentioned earlier, there are two independent variables which are intended to expose the presence of traditional family norms and cultural values. These predictors are the percentage divorced and the percentage one generation family households. Census 2000 is the first census to ask questions about these issues.

In regards to household information among the counties, Census 2000 provides data on the number of households, the number of people in family units, the number of one-person household units, and the number of one, two, three, and four and over generation family units. Extended families are linked to higher rates of fertility, while nuclear families are linked to lower rates of fertility. I decided the best way to measure the variation among my counties in the number of nuclear families was through the percentage of one generation family households. I expect this variable to be negatively related to fertility.

Given that divorce is such a recent phenomenon in the country, it deserves more clarification. I will provide a short synopsis of the cultural role of marriage and confront the recent intrusion of divorce in present-day Chinese society.

Historically, divorce has been widely condemned in the Chinese culture, and much of this has to do with the purpose marriage serves in this society. Rather than being a personal affair of the husband and wife, marriage is an impersonal affair of the families of the husband and wife. "It was stated in the section on marriage definition in the *Book of Rites*, one of the five Chinese classics on Confucian ethical codes, that marriage connects the two sexes for serving upwards the ancestral shrine and continuing downward the descent of the family line" (Zeng et al. 2002: 409). Given this familial purpose, traditional Chinese marriages are arranged by the couple's parents, and are rarely based on love.

Up until the late 1980s, divorce was not only discouraged by the people, it was opposed by the government (Zeng et al. 2002). Chinese citizens are monitored closely by authorities and the same system which regulates fertility also regulates divorce. A couple just over a decade ago faced many obstacles on the road to divorce. It was a multi-step process, which began by submitting a request to their work unit head. The work unit head would then likely disagree and attempt to get them to reconcile. If one or both were adamant, their request was occasionally deferred to the civil affairs department, which would also try to bring about a reconciliation. At this point, family meetings were often called, wherein the couple was once more encouraged to compromise. If the request was deemed rational, the civil affairs department had the option of awarding the divorce. If not, the case would be sent to a local court. The court would follow a similar procedure in trying to dissuade them from getting a divorce. Now at the top of the chain of command, the request would either be granted or turn downed by the court. Countless divorce-seeking individuals gave up at some point in this process, or, after loosing in court, were unwilling to go through the process again.

Socioeconomic development has served to dilute the Confucian ideology over the last twenty years in China. As the society has become more tolerant of divorce, so have the authorities (Zeng et al. 2002). The divorce rate rose from 2.01 per 1000 married couples in 1982 to 3.13 in 1995.

Many socioeconomic variables that have contributed to lower fertility in China have also contributed to higher divorce rates. For instance, the Zeng et al. study (2002) which evaluated the relationship between different socio-demographic indicators and divorce, found urbanization to be positively related to divorce. In this new atmosphere of acceptance, the traditional tenets which once bound couples together sometimes break them apart. "Arranged marriage has a risk of divorce that is about 2.6 times as high as that of non-arranged ones" (Zeng et al. 2002: 422). And "the risk of divorce of women who have three or more daughters without a son was more than twice as high as that of those women who have three or more children with at least one son" (Zeng et al. 2002: 422).

While divorce is still relatively uncommon throughout China today, I believe it is an important variable to include in this thesis particularly since it has not been included in previous analyses of Chinese fertility. Moreover, since I am trying to capture the influence of modernization on fertility, it is theoretically consistent to include divorce as one of my independent variables. The variable will be measured as the percentage of the total population aged 15 and over that is divorced. There are more precise measures of divorce, such as the general divorce rate, which is the number of divorces that took place in year *z* divided by the mid-year total number of married couples. Another reliable divorce statistic is the divorce rate per woman, i.e., the number of divorces for every 1000 married women over 15 years of age. Regretfully my data only permit me to develop this more limited measure of divorce. But since I am really only concerned with the variation among the counties in divorce and since I suspect that the variation in my measure of divorce and the variation in the other divorce rates will be nearly the same, I am comfortable using it in my models. Percentage divorced is expected to have a negative relationship with fertility.

The tenth and final variable is percentage of the total population that is in one of China's 55 minority nationalities. Cultural differences and departures in policy enforcement give rise to higher fertility among minority groups. Therefore, the percentage of minority population is likely to be a key variable in this study, and will be included in many of the models.

Methods

In this thesis, I will estimate a series of Ordinary Least Squares multiple regression models. But there are many preliminary steps I will take beforehand.

First, I will attend to the normality assumption by evaluating the distributions of my dependent and independent variables. While all of the variables do not have to be normally distributed, seriously skewed Y and/or X distributions can produce error distributions which violate the normality assumption. The skewness value in a normal distribution is 0. If skewness surpasses 0.8 in absolute value, then I will know the distribution of the variable is likely skewed. I will also appraise skewness by comparing the values of the mean and median, which are the same in a normal distribution. If the mean of the variable has a lower value than the median, then the distribution is skewed left. If the mean has a higher value, it is skewed right.

I expect that some of my variables will be greatly skewed and this may become a serious impediment to my analysis. Therefore, these problematic variables will be modified using power transformation, a group of straightforward transformations: Y^q or $q > 0$; log Y or $q = 0$; $-(Y^q)$ or $q < 0$. As Hamilton (1992: 18) so clearly notes, different values of q modify the shape of the distribution, such that "powers greater than 1 ($q > 1$) shift the weight to the upper tail of the distribution and thereby *reduce negative skew* and powers less than 1 ($q < 1$) pull in the upper tail and may therefore *reduce positive skew*." As a time saver, rather than exhausting the ladder of powers, people routinely use logarithms when their variables are positively skewed. Under the guidance of my committee chair, Dr. Dudley Poston, I will actually consider several types of transformations and choose the one that best suits each variable.

In STATA, the command **ladder** will employ the most appropriate transformation to convert the variable into a normally distributed variable. The **ladder** command presents the results of the STATA test known as **sktest** for each transformation, which conducts three tests of normality. This includes one test based on skewness and one based on kurtosis, then a joint chi^2 test that is based on both (Poston 2004). In this situation we hope to accept the null hypothesis of normality, and typically, **sktest** would permit our doing this. There are too many observations in my data-set, however, for this command to work properly because chi^2's cannot be estimated. The 2,873 counties will result in an enormous chi^2 for any of the transformations, forcing us to reject rather than accept the null hypothesis of normality. Therefore I will employ the command **gladder**, which considers the same transformations as **ladder**, but displays the results graphically. This will enable us to select the best transformation for each of my variables. Certainly, **sktest** is preferred over visual inspection, but with such a large number of observations, this is the approach I must take.

Many of the independent variables in my analysis are collinear, such as the percentage urban and the percentage nonagricultural. People who are living in urban areas tend to be employed in occupations outside of agriculture. If multicollinearity becomes too excessive, my parameter estimates will be inaccurate. Thus, several measures will be taken to address this impending problem. First, I will examine the zero-order correlations for each pair of the ten variables. I will look at the strength of the correlations. Any correlations above 0.5 or 0.6 indicate possible collinearity problems. Since zero-order correlations have a tendency to obscure certain problematical issues, I will next examine scatterplots of each bivariate relationship. In my examination, I will be looking for heteroscedasticity, curvlinearity, and collinearity troubles. I may then unite predictors that are collinear with each other into indexes, providing that it makes theoretical sense to do so.

Third, I will utilize diagnostic methods, examining tolerance values of the questionable X variables, to determine if multicollinearity is too extreme. In STATA, a good way of doing this is to "quietly" estimate the regression equation, and then use the "vif" command. I will be concerned if the tolerance value for any of my X variables is under 0.35 (i.e. 1/tolerance = VIF), and regression equations with these types of problems will not be estimated. On account of collinearity issues, I plan to split my analysis up into several different models. Each model will be composed of independent variables that have correlations under 0.5 or 0.6. This conservative approach will probably result in models containing no more than three or four independent variables. Consequently, I expect this may require me to estimate several different regression equations.

After estimating my OLS models, I will employ DFBETAs to determine which counties have an overly exceptional impact on the regression coefficients. Lastly, because I expect there may be outlier influence, I will test the validity of my OLS results using Robust Regression. If OLS and robust methods yield consistent estimates of the coefficients, then I will be more assured of my OLS results. If there are any large discrepancies, I will gravitate towards the robust regression results.

In examining my coefficients, I expect to find that counties with a lower percentage of minority population, a higher percentage of non-agriculture population, a higher percentage of urbanization, a higher percentage of white collar workers, a higher percentage of one-generation family households, a higher percentage of divorce, lower percentages of illiteracy, and less gender difference in regards to illiteracy, will have lower total fertility rates.

After performing my county-level fertility analysis, I will create dummy variables for each of the six regions of China. I will go through the same process again, except that I will be examining whether region has a significant impact on fertility. If so, I may then undertake a few region-specific analyses of fertility.

In the next chapter I provide a detailed description of the dependent and independent variables.

CHAPTER IV

DESCRIPTION

Chapter IV provides a statistical description of the dependent variable and ten independent variables to be used in my OLS models. I explore the locations of the extreme values on each variable, looking to cultural, political, and environmental factors for insights. Lastly, I address the specific power transformations performed on each variable, as well as the ensuing means and outermost values. The chapter will therefore be organized around three sections: Dependent Variable, Independent Variables, and Transformed Variables.

According to the 2000 Population Census of China, there are 2,873 Chinese counties or county equivalents. Three of these areas had fertility rates of 0 and were consequently dropped from the dataset. As a result, I will be studying the fertility of the remaining 2,870 Chinese counties or county equivalents. County equivalents are usually cities or districts of cities which the fit the administrative criterion of a county. So bear in mind that in referencing counties, I am also referencing county equivalents. (For sake of simplicity, I will not restate this later in the chapter in my interpretation of each independent variable.)

Dependent Variable

The dependent variable is the total fertility rate (TFR); the TFR is the total number of live births an artificial cohort of 1,000 women would end up having if, as they survived through their 35 years of childbearing, they were exposed to the ASFRs of a county in China in 2000. As indicated by the census data, China's TFR is 1.22. The

TFR among the provinces of China ranges from a low of .67 in Beijing to a high of 2.19 in Guizhou. The mean value of the TFR among the 2,870 counties of China is 1.32 (SD .47). There is considerable variability among the counties in the TFR about its mean value of 1.32, varying from 0.41 in the Xiangyang district of Jiamusi city (Heilongjiang), 0.43 in the Heping district of Tianjin city (Tianjin), and 0.46 in the Mawei district of Fuzhou city (Fujian) to 3.96 in Geji County (Tibet), 4.07 in Jiali County (Tibet), and 5.47 in Baqing County (Tibet) (see TABLE 1). The two extreme scores (i.e. 0.41 and 5.47) produce a range of 5.06. Clearly, the variation of the TFR at the county level is much more dramatic than the variation at the provincial level.

The districts with the three lowest TFRs are found in the provinces of Heilongjiang, Tianjin, and Fujian, respectively located in the northeast, north, and east regions of China. The per capita GDPs of these three provinces are higher than in most provinces of China because they have relatively strong economies and are more advanced in the way of development. Jiamusi City, situated in the heart of Heilonjiang Province, contains the district with the lowest TFR in the nation (Liu 2004). This small city is home to the John Deer Company and is known for its healthy environment and clean air. Jiamusi City is also at the forefront of education, recently launching the Dongze Joy Foreign Language School, which teaches English to 4 - 15 year olds.

The four counties with the highest TFRs are located in the Tibet (Xizang) Autonomous Region, a poor, rural region in northwest China. Baqing County, Tibet, has the highest TFR in the nation. In contrast to Jiamusi City, Baqing County has the highest illiteracy rate and the highest minority percentage among the 2,870 counties of

China. These findings seem to fit the expectations of demographic transition theory (DTT). I will now attend to my independent variables, and in doing so, will provide more contextual information regarding regions, provinces, and counties.

Independent Variables

Three independent variables represent certain aspects of economic development among the counties of China. One is the percentage of the total population that is employed in non-agricultural related work, which has a mean value among the 2,870 counties of China of 27.84 (SD 25.25) (see TABLE 1). This means that 27.84 percent of the population is employed outside the agricultural realm. The variability of percentage non-agriculture is indicated by its wide range (96.84), from lows of 2.07 percent in Jiangda County (Tibet), 2.25 percent in Nanmulin County (Tibet), 2.26 percent in the Hongsibao district of Wuzhong city (Ningxia), and 2.36 percent in Shuicheng County (Guizhou) to highs of 97.76 percent in the Qianjing district of Jiamusi city (Heilongjiang), 97.8 percent in the Lishan district of Anshan city (Liaoning), 97.93 percent in the Shangganling district of Yichun city (Heilongjiang), and 98.91% in the Tiedong district of Anshan city (Liaoning). The two lowest counties are located in the Tibet (Xizang) Autonomous Region, a sprawling, mountainous region (1,201,000 sq km) in southwestern China with a rather small population (2.62 million). Tibet's per capita GDP is low at 3,716 renminbi and its economy is based almost entirely on subsistence agriculture (CPIRC n.d.; CEInet n.d.). The currency conversion is 1.00 United States dollar for every 8.27650 China renminbi. This indicates that, on average, a person living in Tibet makes only $448.97 annually.

The two highest scoring districts on the percentage non-agriculture variable are located in Liaoning Province of northeastern China. Liaoning is spatially smaller (144,900 km) with a population almost sixteen times the size of Tibet's (41.40 million) and a per capita GDP of 9,333 RMB (CEInet n.d.; CPIRC n.d.). Liaoning is one of the country's chief industrial bases and holds four ports, one being the main port of northeastern China (Beauty 1998; Noll 1997). Opportunities for non-agricultural employment are scarce in Tibet and abundant in Liaoning, so these extreme values on this variable seem sensible. Tibet, which contains the two counties with the smallest percentages of non-agricultural workers, also contains the four counties with the highest TFRs. This is indirect support of my hypothesis that percentage non-agriculture will be negatively related to fertility.

The second economic indicator is the percentage of the total population that is urban, which has a mean value of 28.66 (SD 20.96). This means that, among the counties of China, 28.66 percent of the population is residing in an urban area. The variability of percentage urban is extensive, ranging from a minimum of 0.56 percent in Linxia County (Gansu) to a maximum of 100 percent in the Meijiang district of Meizhou city (Guangdong), and in the Dongshi, Zhongshi, and Xishi districts of Bengbu city (Anhui). These extremes result in a range of 99.44. The county with virtually no urban residents is Linxia, located in Gansu Province of northwest China. Like Tibet, Gansu is a large remote province (454,000 sq km) with a rough patchy terrain of grassland, mountains, deserts, and rivers, and given its size, contains a relatively small population (25.62 million) (ChinaTravel n.d.; CEInet n.d.). Linxia County is occupied primarily by the Hui minority nationality, and its environment is quite harsh with cold temperatures and land essentially unreceptive to agricultural production (ChinaPlanner n.d.). Three of the four districts that are exclusively urban are found in Bengbu city in Anhui Province of eastern China. Though also a poor province, Anhui has a population twice the size of Gansu (59.86 million) and is smaller in size (139,400 sq km) (CEInet n.d.). Bengdu city is a port and an influential industrial base, specializing in machinery production (GreatestCities n.d.). Seeing the contextual differences between Linxia County and Bengdu city, their divergent values on the percentage urban variable seem to be appropriate.

The last economic variable is the percentage of the total employed population that is employed in white-collar work. The mean value for percentage white-collar is low at 8.59 (SD 6.49). Thus, managers, administrators, professionals, and associate professionals make up only 8.59 percent of the employed population among the counties of China. The range value for percentage white-collar is 37.25, with a low of 1.37 percent in Dongxiangzu Zizhi County (Gansu) and a high of 38.62 percent in the Xiangyang district of Jiamusi city (Heilongjiang). Jiamusi city holds one of the highest scoring districts in percentage non-agricultural workers, so it is not surprising that it also possesses the district with the largest percentage of white-collar workers. Moreover, Heilongjiang Province, located on the northeastern tip of China has a more advanced economy and a higher per capita GDP (7,544 RMB) than Gansu Province (3,456 RMB) (CEInet n.d.; CPIRC n.d.). The Xiangyang district of Jiamusi city also displays the lowest TFR among the counties of China, which lends support to my theoretical expectations.

Two independent variables reflect the presence of traditional family norms and cultural values. These are the percentage divorced and the percentage one generation family households. The mean value for the percentage one generation variable is 20.93 (SD 7.05), which denotes that 20.93 percent of the households among the 2870 counties of China are one generation family households. The range of percentage one generation is considerable, fluctuating from a low of 3.75 percent in Dongxiangzu Zizhi County (Gansu) to a high of 58.87 percent in the Baoan district of Shenzhen city (Guangdong). In comparison to Gansu's population size of 25.62 million, Guangdong is significantly larger at 86.42 million (CPIRC n.d.). As mentioned earlier, Gansu is an expansive rural province with an erratic landscape, and has one of the lowest per capita GDPs among the provinces of China (3,456 RMB). Located on the southern tip of China, Guangdong is a coastal province with more development and a respectable economy (per capita GDP 11,143). In fact, Guangdong is prosperous enough that it attracts many migrants from other provinces (Wikipedia n.d.). I anticipated that urban areas would have a greater percentage of one generation households, given that urban areas are usually less traditional than rural areas. These divergent findings between Gansu and Guangdong are consistent with my expectations.

The variable percentage divorced has an extremely low mean value of 1.09 (SD 0.79), which is not surprising given the stigma that is attached to divorce in China. This indicates that on average, only 1.09 percent of the population aged 15 and over is divorced among the counties of China. Recall that I could not calculate a divorce rate because the data does not provide information on the number of divorces taking place in 2000. A divorce rate is a more appropriate measurement of this event and yields a different type of figure altogether. However, this more restricted measurement should be adequate for evaluating the variation in divorce among the 2,870 counties.

The range for percentage divorced is 8.07, varying from a low of 0.16 percent in the Hongsibao district of Wuzhong city (Ningxia) to a high of 8.23 percent in Leiwuqi County (Tibet). Wuzhong city possesses the district with the lowest percentage divorced in all of China and is one of the few major cities in the Ningxia Autonomous Hui Region. Located in northwestern China, Ningxia is small in size (66,000 sq km) and population (5.62 million) (CEInet n.d.; Chinastage 1998). Ningxia was integrated into Gansu in 1954, but was disconnected and reestablished as the home of the Muslim Hui minority people in 1958 although they comprise only about 30 percent of the region's population. Ningxia lags behind in development, largely because of a weak infrastructure (Italtrade n.d.). Gravity-fed irrigation from the Yellow River saves this area economically, promoting agricultural activities such as grain production. Even though Ningxia is monetarily poor, it is rich in culture and tradition. In view of the fact that divorce is usually associated with modernization, it seems reasonable that Ningxia has one virtually divorce-free district. What is puzzling, however, are the locations of those counties with high percentages divorced.

Leiwuqi County of the Tibet Autonomous Region has the greatest percentage at 8.23 percent—low by American standards, but considerably high by Chinese. Leiwuqi

County cannot be explained away as an outlier because it does not stand alone; in fact, it is accompanied by many other high scoring Tibetan counties. Tibet has the largest mean on the percentage divorced variable among all the provinces and province equivalences of China. Tibet is thick in custom and culture and the Tibetan minority nationality make up almost 92 percent of its population (Information Office 2004). Tibet was "liberated" by the P.R.C. in 1951 (see "17-Article Agreement"), and it is still spiritually, linguistically, and culturally different from the rest of China. Data aside, it is doubtful that Frank Notestein, the founder of demographic transition theory, would have pegged this underdeveloped rural region as the nation's leader in divorce. On the other hand, it is likely that he would have correctly pinpointed Tibet and Leiwuqi County as high fertility areas. The mean TFR for Tibet is 2.11 and for Leiwuqi County is 3.4; these are scores much higher than the overall mean among all the counties of 1.32.

The other counties with lofty percentages of divorced are Hetian (7.12%), Luopu (6.97%), and Shaya (6.22%), all of which belong to the Xinjiang Uyghur Autonomous Region. Xinjiang trails Tibet among the thirty one provinces and equivalents with the second largest mean on the percentage divorced variable at 2.42%. Xinjiang, situated in northwest China, is a rural region composed of mountains and deserts, and is relatively stagnant in terms of modernization and economic development. In a situation similar to Tibet, Xinjiang is largely inhabited by the Uyghur minority nationality which is known as a "proud, happy, and independent people" (Verber n.d.: 1). Xinjiang's mean TFR is 1.52, and its high divorce counties have mean TFRs of 2.49, 1.95, and 2.06, appreciably higher than the national county mean of 1.32.

These descriptive findings are contradictory in that they support and at the same time, challenge my hypothesis that percentage divorced will be negatively related to fertility. Are these extreme values simply statistical “noise”? This may be the case if the high divorce counties have a small number of people aged 15 and over because it would result in a small denominator, which could produce deceptively high percentages of divorced.

If it is not “noise,” what is causing these unusually high values on this variable? Could it be that modernization is unrelated to divorce in China? Or are certain Chinese minority groups more prone to divorce than the Han majority nationality? These questions are beyond the boundaries of this thesis; nevertheless, I am eager to see how the variable performs in my OLS models predicting fertility.

The next three independent variables are illiteracy percentages, which I will employ to explore the effects of basic educational standing on fertility. The mean value for overall illiteracy is 11.49 (SD 11.19). This indicates that 11.49 percent of the total (male and female) population is illiterate among the 2,870 counties of China. The variable ranges from a low of 0.55 percent in Beiliu city (Guangxi) to a high of 86.22 percent in Baqing County (Tibet). It is not surprising that the high illiteracy county is located in Tibet given the rural and impoverished conditions of this region. Conversely, the minimum extreme is strange and deserves a little more explanation. Beiliu city has almost no illiteracy and it is located in the Guangxi Zhuang Autonomous Region, which borders Guangdong on the south-central tip of China. Guangxi is a mountainous, rural region with a weak infrastructure and a poor economy (Noll 1997; CPIRC n.d.). The region's population numbers 44.89 million and the per capita GDP is low at 4,076 RMB. According to demographic transition theory, Guangxi should contain a county with one of the highest scores on percentage illiterate, not a county with the lowest. I believe the answer to this enigma lies in the size of Beiliu city. Although Beiliu city consists of 23 towns, its entire population is a miniscule 1,186 thousand people (Shuimin 2001). So in all likelihood, its small percentage on the illiteracy variable is just statistical "noise".

As anticipated, a greater percentage of females are illiterate than males. The mean value for male illiteracy is 6.76 (SD 8.55), fluctuating from a low of 0.24 percent in Beiliu city (Guangxi) to a high of 78.16 percent in Baqing County (Tibet). The mean value of percentage female illiteracy is 16.47 (SD 14.27), which is more than twice the mean value of male illiteracy. The range value for female illiteracy is sizeable at 93.44, with a minimum of 0.89 percent in Beiliu city (Guangxi) and a maximum of 94.33 percent in Baqing County (Tibet). The extreme values of male, female, and combined illiteracy occur in the same locations and are explained in the preceding paragraph. While it is bewildering that the lowest score on percentage illiteracy occurs in Beiliu city, Guanxi, the locations of the other low scores on this variable seem sensible. These lows occur in cities and counties of Shanxi, Heilongjiang, and Guangdong.

In China, it is traditionally the case that sons are provided with more educational opportunities than daughters. Daughters are often times expected to forego prospects for personal achievement, since a greater value is bestowed upon sons. Rather than working toward their own aspirations, many daughters often seek employment in the city to help pay for their brothers' education. Given that educational inequality promotes social inequality, I felt that it was important to measure this phenomenon and explore its relationship with Chinese fertility. The percentage female illiteracy divided by the percentage male illiteracy yields a ratio which serves a measurement of female status. This independent variable is interpreted as follows: a value of 1.0 indicates that there is an equal amount of male and female illiteracy, a value of greater than 1.0 indicates that there is more female illiteracy, and a value of less than 1.0 indicates that there is more male illiteracy. It can be inferred that locations with a value approaching 1.0 have a greater degree of equality among the sexes, since both are receiving a similar amount of rudimentary education.

The mean value of female status is 3.16 (SD 1.41), varying from a low of 0.94 in Kangbao County (Hebei) to a high of 16.49 in Pingyuan County (Guangdong). Kangbao County, found in the northern province of Hebei, has the lowest female status score, suggesting that there is a substantial amount of egalitarianism between men and women in this area. In fact, its value on this variable shows that there is more illiteracy among males than females. Hebei has a population of 67.44 million and is in close proximity to two of China's major metropolitan areas—Beijing and Tianjin (CPIRC n.d.; Invest-inchina 2002). This geographical position has benefited Hebei in many ways. Interaction with these neigboring cities has enhanced Hebei's infrastructure and overall development, making it one of the best performers in the government's economic reform effort. Thus, it seems fitting for Hebei to have a high degree of educational equality among the sexes.

The maximum score on the female status variable occurred in Pingyan County, Guangdong. Guangdong is a largely urban and economically advanced province, so it is rather unexpected for it to contain the county with the greatest degree of inequality. Guangdong's overall illiteracy level is much lower than the national county average at 6.63 percent (national county average = 11.69%), so perhaps when both sexes are widely educated, you see more gender difference than you do in places where educational opportunities are scarce. In rural areas, where educational opportunties are limited, there may be less educational inequality because neither sex has sufficient access. I must emphasize that this is all purely speculative.

The last independent variable is percentage minority, which will be very important to my study since minorities tend to have higher fertility rates than the Han majority nationality. There is substantial variability in percentage minority about its mean of 16.14 (SD 28.97), varying from 0 percent in Lin, Shilou, Qingjian, and Zizhou Counties of Shaanxi to 99.78 percent in Angren County of Tibet. This yields a gigantic range of 99.78. Minorities make up over 99.5 percent of the population in ten counties of China and all of these high counties are located in Tibet. This was to be expected, since the Tibetan minority nationality is the overwhelming majority group in Tibet. Minorities comprise less than 0.02 percent of the population in fifty two counties (or county equivalents) of the country. These extremely low minority locations are found in the provinces of Hebei, Shanxi, Jiangxi, Shandong, Henan, Hubei, Guangdong, Sichuan, and Shaanxi.

Transformed Variables

As expected, most of the variables were severely skewed and in need of transformation. For each variable, I produced graphical representations of eight different power transformations (i.e. cubic, square, square root, log, 1/square root, inverse, 1/square, and 1/cubic). Power transformations are "a way to pull in outliers and make skewed distributions more symmetrical, perhaps more normal, and hence easier to analyze" (Hamilton 1992: 148). My thesis chair, Dr. Dudley Poston, assisted me in this process and after we viewed the various transformations, the transformations most closely approximating normal distributions were chosen. I did not use statistical tests for normality because my number of observations was so large, 2,870. Formal statistical tests almost always tended to reject the null hypothesis of normality. We relied therefore on simple visual inspection of the histograms. The dependent variable and nine independent variables were transformed. One independent variable, the percentage of one generation family households, did not necessitate transformation because its skewness value was less than 0.8 at 0.77.

I transformed most of the variables, specifically the TFR, percentage non-agriculture, percentage white-collar, percentage divorced, percentage illiteracy, percentage male illiteracy, percentage female illiteracy, female status, and percentage minority using base 10 logarithms. It makes no difference if one uses natural logarithms or base 10 logarithms. The base 10 logarithm of a variable is the power to which 10 must be raised to yield the value of the variable, whereas the natural logarithm is the power to which "e" must be raised to yield the value of the variable. In some demographic research in which variables are transformed, the natural logarithm is often used instead of the base 10 logarithm. It actually makes no difference with regard to the transformation. Both logarithms have identical effects on the shape of the transformed variable (Hamilton 1992: 17). I will now provide a brief statistical overview of these transformed variables.

The skewness of the dependent variable or TFR, dropped from 1.52 to 0.16 after I transformed it by taking the base 10 log of the TFR. The TFR was positively skewed and the log transformation drew in high values on this variable, making its distribution more normal. The mean of the log of TFR is 0.29 (SD 0.33), ranging from a minimum of -0.89 to a maximum of 1.70.

In its raw form the independent variable, percentage non-agriculture, had a skew of 1.36. The log transformation reduced high outliers, yielding a new skew of 0.31. The mean of the log of percentage non-agriculture is 2.97 percent (SD .84), varying from a low of 0.73 percent to a high of 4.59 percent.

For the independent variable, percentage urban, the square root transformation (q = .5) worked best. While square roots is less powerful than the log transformation, it also brings in extreme outliers. The skewness of percentage urban fell from 1.36 to 0.67 after making this conversion. The mean of the square root of percentage urban is 5.03 percent (SD 1.82), ranging from 0.75 percent to 10.0 percent.

The high skew of percentage white-collar was considerably reduced by using the log transformation, declining from 1.82 to 0.63. The mean of the log of percentage white-collar is 1.94 percent (SD 0.62), with a minimum value of 0.32 percent and a maximum value of 3.65 percent.

Percentage divorced was one of the most abnormally distributed independent variables, possessing a positive skew of 2.58. Once again, of the eight different power transformations, the base 10 logarithm was the best choice, lowering its skew to 0.55. The mean of the log of percentage divorced is -0.11 percent (SD 0.58), fluctuating from -1.83 percent to 2.11 percent.

All three of the illiteracy variables had very poor distributions with severe positive skew and for each of them the base 10 logarithm produced the most normal distribution. The mean of the log of percentage illiteracy is 2.17 percent (SD 0.68), with extreme values of -0.60 percent and 4.46 percent. The mean of the log of percentage male illiteracy is 1.50 percent (SD 0.86), with outermost values of -1.43 percent and 4.36 percent. The mean of the log of percentage female illiteracy is 2.57 percent (SD 0.64), with furthermost values of -0.12 percent and 4.55 percent.

The independent variable, female status, had a positive skew of 2.16 and an extremely abnormal distribution. After evaluating the graphical representations of the eight different transformations, we again chose to use the base 10 logarithm in generating our new variable. The mean of the log of female status is 1.07 (SD 0.39), varying from a low of -0.06 to a high of 2.80. Skewness dropped to 0.37 after making this adjustment.

While most of the independent variables responded well to several of the transformations, the distribution of the percentage minority variable was only improved by the base 10 logarithm. This made our task easier because the choice was an obvious one. The mean of the log of percentage minority is 0.61 percent (SD 2.38), with outermost values of -4.61 percent and 4.60 percent. In making this distribution more symmetric, the skew of percentage minority fell from 1.83 to 0.15.

China has finally become a lowest-low fertility population; however, fertility differs markedly among the nation's regions, provinces and counties. This statistical description has illustrated how drastic variation amid social and economic factors may contribute to these fertility fluctuations. The following chapter presents the results and interpretations of my OLS regression models. Two duplicate analyses will be discussed: one that does not include region and one that does.

CHAPTER V

REGRESSION ANALYSES OF THE TOTAL FERTILITY RATE

In this chapter, several Ordinary Least Squares regression equations are estimated to test the strength and capacity of demographic transition theory (DTT) in predicting Chinese fertility. First, I discuss at length the diagnostic strategies that helped in preventing statistical problems and developing models later in the study. Next, I present the results of three multiple regression models across the 2,432 counties of China. The number of counties has fallen from 2,870 to 2,432. This is because 438 counties had missing data on one or more of the independent variables. So to maintain consistency, I dropped all 438 and used the same number of counties in all equations for the duration of the analysis. Finally, I present the results of the same three models, while controlling for the effects of region.

Diagnostic Strategies

To ensure that my OLS models were statistically sound, I used several diagnostics prior to the analysis. Initially, I assessed the distributions of the dependent and independent variables, and I found all but one of the variables to be skewed (i.e. skewness was greater than 0.8 in absolute value). Heavily skewed Y and/or X distributions can produce nonnormal error terms, which thus violate the normality assumption. Consequently I decided to conduct power transformations on these variables to draw in the positive outliers; most were transformed using the base 10 logarithm. Please note that when referring to variables over the course of this chapter, I will refer to their transformed versions. The only variable that did not require transformation is percentage households with one generation.

Next, I examined the zero-order correlations of the independent variables with the TFR and with each other (see **TABLE 2**). All of the independent variables were associated with the TFR in the direction my hypotheses predicted, except for female status (r = -0.115). This is not overly shocking, since the outermost values on that variable had already alerted me to potential problems. The high county on that variable is located in the relatively modernized province of Guangdong, which would go against theoretical expectations. As reported in **TABLE 2**, the three urbanization and economic variables have moderate to moderately high correlations with fertility: percentage labor force white collar, -0.455; percentage population urban, -0.560; and percentage population nonagriculture, -0.568. Both of the variables reflect traditional family norms and cultural values and show the hypothesized negative association with fertility: percentage households one generation, -0.422 and percentage 15+ divorced, -0.087. After discovering the strange locations of the high percentage divorced counties, I was concerned that divorce may have a positive rather than a negative relationship with Chinese fertility. Even though the relationship between fertility and divorce is extremely weak, as illustrated by the correlation of -0.087, I was relieved to see that it was negative as I had initially imagined it would be. The three variables which assess illiteracy have modest correlations with fertility: percentage illiterate, 0.439; percentage male illiterate, 0.394; and percentage female illiterate, 0.441. A little perplexing is the very fairly small correlation between percentage minority and fertility of 0.232. Among the counties of

Sichuan Province, this correlation was much stronger at 0.542, and I had expected to see an even higher degree of association in this analysis. Part of the reason for the diluted association is that I am using the transformed variables in this thesis, whereas I used the raw variables in the previous study.

Several of my independent variables were highly correlated with one another, which signified collinearity problems. The following independent variables have correlations with each other above 0.5: percentage urban and percentage households one generation, 0.538; percentage male illiterate and female status, -0.665; percentage urban and percentage white collar, 0.754; percentage nonagriculture and percentage white collar, 0.838; percentage nonagriculture and percentage urban, 0.845; percentage male illiterate and percentage female illiterate, 0.878; percentage illiterate and percentage male illiterate, 0.941; and percentage illiterate and percentage female illiterate 0.987. These high correlations were expected given that most of these variables are clearly related to one another (e.g. people who are engaged in nonagricultural employment are more likely to have white collar occupations).

Scatterplots of these bivariate relationships confirmed my suspicions regarding the highly correlated independent variables. For instance, the graph of percentage illiterate with percentage female illiterate showed a tight, positive, straight line pattern (indicating that almost all of the variance in one is accounted for by the other). Scatterplots illustrating the bivariate relationships of the TFR with each of the ten independent variables (see **FIGURES 1-10**) were especially telling. Positive and negative associations are easily identified in each of the plots by the directional flow of the observations.

After these preliminary checks, I was ready to split my analysis up into different models. This process was also performed in steps and with particular questions in mind. The zero-order correlations had alerted me to the highly associated illiteracy variables. All three showed the expected positive relationships with fertility. But since they could not be estimated together in the same OLS models, I discarded two and retained only one for my analysis. My other concern was regarding the female status variable, which had a relationship with fertility that ran counter to my hypothesis. I anticipated that female status, measuring educational inequality between the sexes, would be positively associated with fertility, but instead the association was negative. Typically, when correlations are negative, effects are negative, but I wanted to be sure before taking female status out of my analysis.

Extensive pairing and exploration with different regression models helped me to determine that female illiteracy had a greater impact on fertility than the two other illiteracy variables. Predominantly, the female status variable had negligible impact on fertility. And when the effects were significant, they contradicted the assumptions of demographic transition theory.

Obviously, results are not always consistent with our theoretical expectations and when they deviate, it is inappropriate for a researcher to remove the offending variables from the analysis altogether. Along these lines, someone might say that the disruptive effect of female status is an important finding and one to be explored throughout the study. There a few factors, however, that would prevent me from taking this stance regarding the female status variable. First of all, ratios are known to cause problems in quantitative studies like this one; numerator and denominator effects often distort measurement outcomes. Accordingly, the strange association of female status with fertility could be a methodological issue. My second source of skepticism is the two decades worth of research verifying this negative relationship between female status and fertility in China. This furthers my belief that this is simply a statistical snag, not a reflection of reality. Third, employing illiteracy percentages in my development of the female status variable may have been an unsuitable choice of action, given the high degree of literacy that exists throughout China today. Census 2000 provides an abundance of data on educational attainment, which may have been used to achieve a more accurate measurement of inequality among the sexes. Perhaps I should have measured female status by taking the ratio of the percentage of females completing high school to the percentage of males completing high school. Further research will need to consider some of these options.

For these reasons, before developing my definitive OLS fertility models, I dropped percentage illiteracy, percentage male illiteracy, and female status from the pool of independent variables. I now had seven predictors remaining.

To guard against collinearity problems, I decided each model would only be composed of independent variables whose correlations were under 0.60. I also chose to eliminate models that produced tolerance values below 0.40 for any of the variables. In taking this conservative approach to model building, however, I had no tolerance problems, and so elimination was never an issue.

The three economic development variables, percentage urban, percentage nonagriculture, and percentage white collar, were highly associated with one another (Pearsonian r's > .75), so I could not include them together in the same model. The other four variables, percentage one generation, percentage divorced, percentage minority, and percentage female illiterate, had inconsequential relationships with one another and the economic variables, indicating that they could be simultaneously estimated.

In view of the three highly correlated economic development variables, I decided to divide my analysis into three distinct OLS models (see **TABLE 3**). Each model contains five independent variables and the TFR. The four cultural, minority, and illiteracy independent variables are the same in all three models; the single economic variable is what sets the models apart from one another.

As noted previously, I expect the cultural and economic development variables to have negative effects on fertility, while I expect the minority and illiteracy variables to have positive effects on fertility. **TABLE 3** reports the effects of these independent variables on the TFR for all three multiple regression models. Standardized estimates are listed in parentheses under the coefficients and significance values are specified with asterisks (see key beneath **TABLE 3**). Recall that I have transformed all but one of my variables, which makes the full interpretation of their respective coefficients somewhat unconventional and unwieldy. So in the interpretations of the regression coefficients, I will not state them precisely using the transformations in my interpretations. For example, I will not be interpreting my regression coefficients in the following way: For every one unit increase in the square-root of percentage urban, there is an average .060 decrease in the $\log_{10}$ of the TFR, provided that the values on the other independent variables remain constant. Instead, for the sake of brevity, I will not engage in such detailed explanations and will address my estimates as simple positive or negative associations. I will be interpreting the above association as follows: Percentage urban is negatively and significantly associated with the TFR, controlling for the other variables in the equation. Also, in many of my interpretations, I will not be adding the phrase "controlling for the other variables in the equation." Each interpretation, however, obviously carries with it the above proviso.

Models 1-3

Model 1 examines the impact of five independent variables on the TFR. The economic development variable in this regression analysis is the percentage of the population that is urban, and is expected to have a negative relationship with the TFR. That is to say, the higher the percentage of urban residents in a county, the lower the fertility. The two variables intended to measure traditional family norms and cultural values, the percentage of households that are one generation and the percentage of the population (15+) that is divorced, are also both expected to be negatively associated with the TFR. Ethnic differences and more flexible policy requirements should lead to higher fertility rates among minority groups; thus, I anticipate the minority variable to have a positive relationship with the TFR. Specifically, I hypothesize that counties with a high percentage of minority population will have high rates of fertility. The last independent variable is the percentage of the female population (15+) that is illiterate, and it is expected to be positively associated with the TFR.

The coefficients in Model 1 all show significant effects on fertility in the hypothesized direction: percentage urban (-), percentage one generation (-), percentage divorced (-), percentage minority (+), and percentage female illiterate (+). Each of the five coefficients is significant at $p < 0.005$. The R^2 (adjusted) statistic is 0.4212, meaning that a little over 42 percent of the variance in fertility is accounted for by these three X variables. The F is a high 354.77, $p = 0.000$, indicating that it is improbable that R^2 in the population is zero. In my evaluation of the standardized coefficients, I see that percentage urban has the strongest influence over fertility among these five variables, $b^* = -.371$, followed by percentage female illiterate, percentage minority, percentage one generation, and percentage divorced. While the effect of percentage divorced on the TFR is comparatively weak, I was excited to see that the relationship was significant and in the hypothesized negative direction. This model supports the notions of demographic transition theory.

Model 2 uses four of the same independent variables, percentage one generation, percentage divorced, percentage minority, and percentage female illiterate to predict the TFR. The difference between Model 1 and Model 2 is that percentage urban has been replaced by percentage nonagriculture, which is also expected to have a negative effect on fertility. Once again, all of the coefficients' signs are in the expected direction and all are significant at $p < 0.005$, with the exception of the percentage divorced coefficient, which has the hypothesized effect but is not significant at 0.05. The following are some general interpretations of the coefficients in Model 2. As the percentage of nonagricultural workers increases, the TFR decreases, controlling for the other variables in the equation. For every increase in the percentage of one generation households (or nuclear family households), there is a corresponding decrease in the TFR. The greater the percentage of minority population, the higher the TFR, all things equal. And finally, with every increase in the percentage of female illiterate, there is a subsequent increase in the TFR.

In Model 2, the percentage of the population employed in nonagricultural work has the largest impact on the TFR, b* = -.405, trailed by percentage female illiterate, percentage one generation, and percentage minority. The R^2 (adjusted) is slightly larger in Model 2 than in Model 1 at 0.4335, indicating that the X variables explain just over 43 percent of the variation in the TFR among the counties of China. The null hypothesis that all five b coefficients are zero can be rejected, as F = 373.00, p = 0.000. The two economic development variables, percentage urban and percentage nonagriculture, have the largest relative effects on fertility in Models 1 and 2, lending further support to demographic transition theory.

Model 3 presents the results of another multiple regression analysis of the TFR on the same four cultural, minority, and illiteracy variables, but this time the economic development variable is percentage white collar workers. This new variable is expected to have a negative impact on the TFR. The coefficients reveal that all of the independent variables have significant effects on fertility in the hypothesized direction: percentage white collar (-), percentage one generation (-), percentage divorced (-), percentage minority (+), and percentage female illiterate (+). The coefficient of percentage divorced is significant at $p < 0.05$, and the other four coefficients are significant at $p < 0.005$. Considering the standardized coefficients, percentage female illiterate has the strongest effect on fertility, $b^* = 0.281$, followed by percentage white collar, percentage one generation, percentage minority, and percentage divorced. Given that there is such a small percentage of white collar workers among the counties of China, it seems sensible that the effect of this economic development variable is less than the effects of those in the previous two models. The X variables in Model 3 account for 38 percent of the variation in the TFR among the 2,432 counties of China and the H_o may be rejected; $F = 299.83$ and $p = 0.000$.

An indication of robustness is that the adjusted coefficients of determination (R^2) do not vary much among the three different models, ranging from 0.3807 to 0.4335. I examined tolerance values to determine if multicollinearity was too extreme in any of the three regression equations. None of the tolerance values dropped below 0.40 for any of the X variables; thus, collinearity is not a problem. Model 3 has the lowest tolerance value of 0.59 on the percentage nonagriculture variable, which means that 59 percent of the variance in percentage nonagriculture is unrelated to the other four independent variables. Such high tolerance values were expected, since I only used predictors with inter correlations of 0.60 or less.

I re-estimated these three models using robust regression equations and found the outcomes to be very comparable to my OLS outcomes. The size and significance of the coefficients remained constant, bolstering my confidence in the OLS results.

These results support my general hypothesis that, among the 2,432 counties of China, the higher the levels of social and economic development, the lower the fertility.

Models 4-6: Regional Controls

As I mentioned previously, I felt that it was important to test demographic transition theory while controlling for region, since theoretically, DTT is expected to work in any given setting. Therefore, in this next part of my OLS analysis, I estimate the same three models, but add to each of them dummy variables for the six regions of China. Since I had to choose one region for the reference group, I first examined the mean TFRs among the counties in each of the regions. The results are as follows. The North region contains the provinces of Beijing, Tianjin, Hebei, Shanxi, and Inner Mongolia and has a mean TFR of 1.31. The Northeast region consists of Liaoning, Jilin, and Heilongjiang and has a mean TFR of 0.95. The East region includes Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, and Shandong and has a mean TFR of 1.24. The Central South region holds the provinces of Henan, Hubei, Hunan, Guangdong, Guangxi, and Hainan and has an average TFR of 1.39. The Southwest region includes Chongqing, Sichuan, Guizhou, and Yunnan and its mean TFR is 1.71. And finally, the Northwest region contains the provinces of Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xingjiang and has a mean TFR of 1.51. Based on these regional TFRs, I chose the Northeast region as my reference group with the lowest overall TFR of 0.95.

Models 4-6 (see **TABLE 4**) present the results of three multiple regression analyses of the TFR on the various socioeconomic and cultural variables, while controlling for region. As stated above, the DTT variables that were used in Models 1-3 are exactly the same in Models 4-6, the only difference being the introduction in Models 4-6 of the region dummies.

In Model 4, four of the five DTT coefficients are statistically significant at $p < 0.005$ and are signed in the expected direction. All things equal, percentage urban is negatively related to the TFR, percentage one generation is negatively associated with the TFR, percentage minority is positively related to the TFR, and percentage female illiterate is positively tied to the TFR. The coefficient of percentage divorced is neither significant nor shows the predicted effect. The region coefficients are all positively and significantly related to the TFR, but they are included in the model only as controls, not substantive predictors. The coefficients of these region variables are all positive. This means that the fertility rates of the counties in the five regions are all higher than that of the reference region, the Northeast, controlling for the effects of the substantive DTT variables. This means that even after considering the DTT variables included in the equations, there remain differences in fertility among the regions.

DTT seems to have fared well even under the regional controls so far, as the DTT coefficients in Model 4 are very comparable to those in Model 1. I will now track their changes (please note that when I say increase or decrease, I am referring to the absolute value of the coefficient). Percentage urban decreased from -0.060 in Model 1 to -0.051 in Model 4; percentage one generation increased from -0.007 to -0.008;

percentage minority stayed the same at 0.020; and percentage female illiterate decreased from 0.120 in Model 1 to 0.086 in Model 4. The R^2 (adjusted) in Model 4 is moderately high at 0.4953, meaning that the variables in this model account for nearly half of the variation in the TFR. This is just a little higher than the R^2 (adjusted) in Model 1 of 0.4212 because differences among the regions are now taken into account.

Model 5 is a duplication of Model 2, except that I have incorporated the region dummy variables. Once again, excluding percentage divorced, the coefficients of the DTT variables show the hypothesized effect on the TFR and are statistically significant at $p < 0.005$. Even when controlling for region, percentage nonagriculture and percentage one generation are both negatively related to the TFR, while percentage minority and percentage female illiterate are both positively associated with the TFR. The size of the coefficients in Model 2 and Model 5 are similar. Percentage nonagriculture decreased from -0.175 to -0.139, percentage one generation increased from -0.008 to -0.010, percentage minority decreased from 0.021 to 0.020, and percentage female illiterate decreased from 0.098 in Model 2 to 0.076 in Model 5. This speaks well for DTT in predicting Chinese fertility. The R^2 (adjusted) is augmented slightly from 0.4335 in Model 2 to 0.4942 in Model 5.

Model 6 is estimating the effects of the same DTT variables on fertility as Model 3, while controlling for region. Percentage white collar is negatively and significantly associated with the TFR, controlling for the other variables in the equation, including region. Percentage one generation is negatively and significantly related to the TFR. Percentage minority is positively and significantly associated with the TFR, everything else equal. And percentage female illiterate is positively and significantly related to the TFR, controlling for percentage white collar, percentage one generation, percentage minority, percentage divorce, and region. All the coefficients are significant at $p < 0.005$. Again, the coefficient of percentage divorced is both insignificant and signed in the wrong direction. The fact that this has occurred in all three models suggests that region is related to divorce and when controlling for region, we dilute the effect of divorce on the TFR. I am not too concerned about percentage divorced, however, since it is not exactly a bona fide DTT variable as are percentage urban, percentage nonagriculture, percentage white collar, and percentage female illiterate. These coefficients barely change from Model 3 to Model 6, which indicates that DTT is reasonably successful at predicting Chinese fertility. The R^2 (adjusted) increases from 0.3807 in Model 3 to 0.4613 in Model 6. The null hypothesis that all the b coefficients are zero can be rejected, as $F = 209.17$ and $p = 0.000$.

I employed robust regression to test the accuracy of the OLS results for Models 4-6 and saw very little change in the performance of the DTT predictors. This reassured me of my original OLS findings.

The fundamental DTT variables in all three models have maintained their statistical significance while controlling for region. In this analysis, I was able to examine two new variables, percentage (15+) divorced and percentage one generation households. While percentage divorced became insignificant after controlling for region, percentage one generation maintained significance and strength throughout the analysis. This illustrates that not only do urbanization and modernization decrease fertility, so does a break with the traditional family household.

The effect of percentage divorced on fertility may be weakened to some extent by the conflicting origins of divorce in present-day China. Zeng and his colleagues (2002) found that urbanization is positively related to divorce, which is compatible with the expectations of demographic transition theory. However, their research also showed that in some instances the risk of divorce is higher among persons with arranged marriages and among women who fail to produce a son by the time they have their third child. Thus it appears that divorce is linked to secularization and lower levels of fertility in some areas of the country, while it is linked to traditional beliefs and higher levels of fertility in other parts of the country. Since I included all 2,432 counties in this phase of my fertility analysis, these contradictory relationships were likely captured in my OLS equations and are reflected in the results. Perhaps when this variable is employed in later region-specific analyses, its effect will be stronger. If in fact this is the case, I can assume it is because these types of inconsistencies were overcome by a more homogenous group of counties (i.e. mainly secularized or mainly traditional).

These OLS results suggest that DTT is valid for predicting and understanding fertility among the counties of China. Even with the nation's legendary family planning policy, other factors indeed play a role in the fertility outcomes of the people of China.

In the next chapter, I will present the results of two region-specific fertility analyses.

CHAPTER VI

REGION-SPECIFIC REGRESSION ANALYSES OF THE TOTAL FERTILITY RATE

In the last chapter, it was shown that demographic transition variables are quite successful predicting fertility among the 2,432 counties of China (see **TABLE 3**). Most of these variables maintained consistency in effect and significance, even while controlling for region (see **TABLE 4**). At the same time, it was apparent that region also played a large role in determining the fertility levels among the Chinese counties. Each of the region dummy variables had a positive and statistically significant effect on the TFR in all three models. Given these findings, Dr. Poston and I decided that it would be wise to conduct two region-specific fertility analyses, one for the highest fertility region of China and one for the lowest. This approach would be a more accurate and stringent test of DTT, ultimately bolstering or weakening our confidence in the theory's capacity for predicting Chinese fertility.

Chapter VI thus presents the results of six models. The basic three models (shown in **TABLE 3**) from Chapter V are employed in two region-specific analyses. The first fertility analysis focuses on the Northeast Region, which has the lowest mean TFR among the six regions of China of 0.95. The second analysis involves the Southwest Region, which has the highest mean TFR of 1.71. (Note that the means of these regions are calculated from the counties within these regions, not from the provinces. In other words, though I am shifting to a regional focus, this is still a county-level investigation.) When I discuss the results of Models 7-12, I will compare them to the results of Models 1-3.

Models 7-9

Models 7-9 (see **TABLE 5**) pertain to the counties of the Northeast, the lowest fertility region of China. The Northeast Region contains the three relatively modernized provinces of Liaoning, Jilin, and Heilongjiang. As you, the reader, will soon see, this situation alters the predictive impact of the DTT variables in a unique way.

Model 7 examines the effects of five independent variables on the TFR among the 216 counties of Northeast China. The predictors are the same as in Model 1. The percentage of the population that is urban, the percentage of households that are one generation, and the percentage of the population (15+) that is divorced, are the independent variables expected to have negative relationships with the TFR. In contrast, the percentage of the population that is minority, and the percentage of the female population (15+) that is illiterate are the independent variables anticipated to have positive associations with the TFR.

The coefficients in Model 7 show all of the independent variables as having significant effects on fertility in the hypothesized directions, except for percentage one generation: percentage urban (-), percentage divorced (-), percentage minority (+), and percentage female illiterate (+).

Model 7 demonstrates that even in the Northeast, as the percentage of urban residents in the population increases, the TFR decreases, controlling for the effects of the other independent variables. The significance of this variable has fallen from $p < 0.005$

in Model 1 to $p < 0.05$ in Model 7. But this is due in part to a reduction in the number of observations. In contrast to the countrywide county-level analysis where much of the population is rural, this analysis was performed on a largely urban population. Hence, the influence of this variable was weakened to some extent.

The performance of the two variables intended to expose traditional family norms and cultural values, percentage divorced and percentage one generation, differed markedly in Model 7. Unexpectedly, the most notable statistic in the entire model is the coefficient of percentage divorced, significant at $p < 0.005$. Among the five variables, percentage divorced has the largest standardized impact on the TFR, $b^* = -0.370$, followed by percentage female illiterate, 0.317, percentage minority, 0.264, and percentage urban, -0.226. In Model 1, percentage divorced was also significant at $p < 0.005$, but it exerted the least amount of influence among these same five variables. Percentage urban and percentage divorced are perfect illustrations of how geographical differences may reduce the effect of a particular variable, while inflating the effect of another.

In Model 1, among all the counties of China, the percentage one generation variable did quite well, displaying both a negative and significant effect on the TFR at $p < 0.005$. Conversely, percentage one generation households had neither the hypothesized effect nor a statistically significant impact on the TFR in Model 7. So limiting the analysis to the low fertility region eradicated its influence altogether. Conceivably in the predominately urbanized Northeast Region, most households are nuclear family households, with only a handful of households being composed of extended families. If in fact this is the case, there may not have been sufficient variation for this variable to function properly in the OLS regression equation.

As in Model 1, percentage minority and percentage female illiterate are both positively and significantly related to the TFR in Model 7 at $p < 0.005$. This means that with each increase in the percentage of minority population, there will be a corresponding increase in the TFR among the 216 counties of Northeast China, all else equal. Additionally, it may be said that with each decrease in the percentage of females (15+) that are illiterate, there will be a subsequent decrease in the TFR, holding constant the effects of the other independent variables on the TFR.

The R^2 (adjusted) statistic indicates that the independent variables in Model 7 explain almost 45% of the variation in the TFR among the counties of the Northeast. The null hypothesis that all five predictors are zero is easily rejected, as $F = 35.97$ and $p = 0.000$.

Model 8 estimates the effects of percentage nonagriculture, percentage one generation, percentage divorced, percentage minority, and percentage female illiterate on the TFR. The predictors are identical to those in Model 2.

As occurred in Model 7, the one independent variable in Model 8 that operates poorly on the TFR is the percentage of one generation households; it is both insignificant and signed in the wrong direction. The other four variables show significant effects in the hypothesized direction.

The percentage of the population employed in non-agricultural related work is negatively associated with the TFR among the counties of the Northeast and is significant at $p < 0.05$. The coefficient of this variable was slightly more significant in Model 2 at $p < 0.005$. This change is reminiscent of that which occurred with the percentage urban variable in Model 7. It is likely that the effects of both of these variables were reduced in part because of the developed living conditions in the Northeast. Much of the population is urbanized and engaged in non-agricultural employment, so the measures were not afforded the variability they had in the nationwide county-level analysis.

In Model 2, the coefficient of percentage divorced was insignificant, but in Model 8, it is very significant at $p < 0.005$. With each increase in the percentage of the population (15+) that is divorced among the counties of the Northeast, there is a resulting decrease in the TFR, all else equal.

As expected, percentage minority and percentage female illiterate have negative relationships with the TFR in Model 8 and their coefficients are both significant at $p < 0.005$.

In my evaluation of the standardized coefficients, I note that percentage female illiterate and percentage divorced have the strongest influence over fertility among these five variables, with respective b*'s of 0.330 and -0.330. These findings are much different from those of Model 2, where percentage nonagriculture had considerably more relative impact than the other four independent variables.

The R^2 (adjusted) statistic in Model 8 is 0.4501, meaning that 45% of the variance in fertility among the counties of the Northeast is accounted for by these five independent variables. The F is 36.20 and $p = 0.000$, indicating that it is unlikely that R^2 in the population is zero.

Model 9 presents the results of the same multiple regression equation shown in Model 3, except of course, that it is specific to the counties of the Northeast. The coefficients show only three of the independent variables as having significant effects on fertility in the hypothesized directions: percentage minority (+), percentage female illiterate (+), and percentage divorced (-).

As expected, percentage white collar has a negative relationship with the TFR, but it is insignificant at all levels. In Model 3, this variable was significant at $p < 0.005$ and as evidenced by the standardized coefficients, it had one of the strongest effects on the TFR among the five independent variables. Among the developed counties of the Northeast, it is apparent that fertility outcomes have little to do with the percentage of the labor force that is employed in white collar occupations. Why is this so? In the nationwide county-level fertility analysis, percentage white collar served as a reliable measure of economic development and modernization since managers, administrators, professionals, and associate professionals are almost always employed in urban areas. But when limiting my analysis to the Northeast Region, I am essentially limiting it to urbanized areas. Therefore, the variable loses its function as an indicator of urbanization and development. Recall that only 5.92% of the labor force in China is employed in these white collar occupations, so the measure itself is very restricted, barring its purpose as an overall indicator of economic development. It seems plausible that this variable was only significant in Model 3 because it is so highly related to other markers of urbanization, such as percentage nonagriculture and percentage urban. Both of these variables were left out of the regression equation for precisely this reason; all three variables (i.e. percentage white collar, percentage nonagriculture, and percentage urban) are highly collinear with one another.

Model 9 also shows percentage one generation to be ineffective; its coefficient is both insignificant and signed in the wrong direction. I have already discussed my reasoning behind this variable's diluted impact on fertility in the Northeast, so I will now proceed with a short discussion of the significant independent variables.

The coefficients of percentage divorced, percentage minority, and percentage female illiterate are all significant at $p < 0.005$. According to the standardized coefficients in Model 9, percentage divorced has the strongest impact on the TFR, $b^* = -0.445$, followed by percentage female illiterate, 0.337, and percentage minority, 0.282. In Model 3, percentage divorced had the smallest effect on fertility among the five predictors and was less significant at $p < 0.05$. The influences of percentage female illiterate and percentage minority on the TFR are also intensified in Model 9, but the differences were less drastic than that of percentage divorced.

The independent variables in Model 9 account for 44% of the variation in the TFR among the 216 counties of the Northeast and the null hypothesis may be rejected; $F = 34.40$ and $p = 0.000$.

The R^2's (adjusted) do not differ much among Models 7-9, ranging from 0.4372 to 0.4501. This indicates that there is robustness about this combination of variables in explaining fertility in the Northeast. I also inspected tolerance values to determine if multicollinearity was a problem in any of the regression equations and found that none of the tolerances dropped below 0.30 for any of the independent variables. While these tolerances are a bit lower than those of Models 1-3, that was to be expected since there are 2,216 fewer observations in this analysis. Furthermore, the counties in the Northeast are more homogenous than those spanning the entire nation. I re-estimated Models 7-9 using robust regression and almost all of the coefficients are less than one standard error from the analogous robust coefficients. Thus, I am confident about my reporting of the OLS results.

Models 7, 8, and 9 suggest that divorce has the largest impact on Chinese fertility in the developed Northeast—a very intriguing discovery! Liaoning, Jilin, and Heilongjiang are all relatively modernized, so the urbanization and economic development variables are having a diminished impact on the TFR among the counties in those provinces. The percentage minority and percentage female illiterate variables have maintained their significance and strength of effect in the three Northeast multiple regression fertility models. This analysis has shown mixed results regarding the power of DTT variables in predicting fertility. I will move now to a discussion of Models 10, 11, and 12, all of which examine the TFR in the Southwest, the high fertility region of China.

Models 10-12

Models 10-12 (see **TABLE 6**) correspond to Models 1-3 and Models 7-12, save that they are examining fertility in the counties of the Southwest Region, whereas the other six models examined fertility in the counties across the nation and in the Northeast

Region. Recall that the mean TFR among the counties of the Southwest is 1.71, the highest among the six regions of China. The Southwest Region contains the provinces of Sichuan, Guizhou, Yunnan, and the Municipality of Chongqing. A comparison of per capita GDP among these provinces gives one an idea of the economic and developmental differences that exist between the Northeast and Southwest Regions. The values of average per capita GDP for the three provinces of the Northeast are 9,333, 5,916, and 7,544 renminbi in Liaoning, Jilin, and Heilongjiang, respectively. These averages are much higher than those among the four provinces of the Southwest of 4,684, 4,339, 2,342, and 4,355 renminbi in Chongqing, Sichuan, Guizhou, and Yunnan, respectively (CPIRC n.d.).

Model 10 is equivalent to Models 1 and 7, except that it is estimating the effects of the five independent variables on fertility among the 415 counties of the Southwest Region. Three of the variables performed very well in Model 10, while two performed very poorly. At any rate, the model has an extremely high R^2 (adjusted) at 0.6361, which is significant as $F = 145.76$ and $p = 0.000$. This indicates that the independent variables in Model 10 account for almost 64% of the variation in the TFR.

Percentage urban and percentage one generation were both expected to be negatively associated with the TFR, but neither of their coefficients are statistically significant in Model 10. The coefficient of percentage urban maintained statistical significance when the regression equation was estimated for the counties of the Northeast region; however, in the counties of the Southwest it appears that percentage urban has little impact on the TFR. Perhaps there are an insufficient number of urban residents in the counties of the Southwest for this variable to work properly as a predictor. The coefficient of percentage one generation households is signed in the wrong direction and is insignificant in Model 10, just as it was in Model 7. The percentage one generation variable performed well in Models 1-3, across all the counties of China, so it is fascinating that it fails as a predictor in both the high and low fertility regions.

Percentage divorced is negatively related with the TFR in Model 10 and is statistically significant at $p < 0.005$. As the percentage of the population (15+) that is divorced increases, the TFR decreases, holding constant the effects of the other independent variables. This variable was also significant at the same level in Models 1 and 7.

The independent variables anticipated to have a positive association with the TFR, percentage minority and percentage female illiterate, show the hypothesized effects and are both significant at $p < 0.005$. These variables also showed positive associations and identical significance levels in Models 1 and 7.

The standardized coefficients reveal that percentage female illiterate has the strongest influence on fertility among these five variables, $b^* = 0.437$, followed by percentage divorced, -0.363, and percentage minority, 0.335.

I will now address the results of Model 11, which will be compared with the results of Models 2 and 8. The coefficients in Model 11 indicate that all of the independent variables have significant effects on the TFR in the hypothesized direction, except for percentage one generation: percentage nonagriculture (-), percentage divorced

(-), percentage minority (+), and percentage female illiterate (+). The R^2 (adjusted) is 0.6432, which indicates that these predictors account for 64% of the variance in the TFR. This is very similar to the value of the R^2 (adjusted) in Model 10.

The standardized coefficients are a helpful source of information when comparing the strength of variables in different models. In Model 11, percentage female illiterate has the most powerful influence on the TFR, as b* = 0.391. It is followed by percentage minority, 0.348, percentage divorced, -0.318, and percentage nonagriculture, -0.147. These results are much different from those in Model 2, where percentage nonagriculture had the strongest impact among these five variables, trailed by percentage female illiterate, percentage one generation, and percentage minority. Model 8 shows the force of these variables in shaping the TFR of the Northeast. In this model, percentage female illiterate and percentage divorce have the greatest effects, followed by percentage minority, and percentage nonagriculture. Percentage one generation is insignificant in Model 11 and Model 8 and percentage divorced was insignificant in Model 2. Clearly, these five variables carry distinctive weights and have altered impacts when predicting fertility outcomes among the counties of the Southwest Region, the Northeast Region, and the entire nation of China.

I will now discuss Model 12, the last OLS model to be presented in this thesis. The variables in Model 12 are identical to those in Models 3 and 9. To remind the reader of the hypothesized relationships regarding the DTT variables, the economic development variable in this regression equation is the percentage of the labor force that is employed in white collar work, and it is expected to have a negative relationship with the TFR. That is to say, the higher the percentage of white collar workers, the lower the fertility. The two variables projected to reflect traditional family norms and cultural values, the percentage of households that are one generation and the percentage of the population (15+) that is divorced, are also both expected to be negatively associated with the TFR. Ethnic distinctions and less restrictive fertility regulations bring about higher fertility rates among minority groups; therefore, I anticipate the minority variable to have a positive relationship with the TFR. And finally, demographers have established that education has a negative impact on fertility, so I employed female illiteracy as an indicator of deficient educational standing. I hypothesize that percentage female illiterate will be positively associated with the TFR. Namely, I expect that counties with a higher percentage of illiterate females will have higher rates of fertility.

Two variables are unsuccessful in Model 12. These are percentage white collar and percentage one generation. The coefficient of percentage white collar is signed in the right direction, but it is not statistically significant. The coefficient of percentage one generation is signed in the wrong direction and it is also insignificant. Bear in mind that that these are the same two variables that failed in the exact same manner in Model 9, the analogous Northeast model. This means that it makes no difference whether we are predicting fertility in the highly developed, low fertility region of the Northeast or the rural, underdeveloped, high fertility region of the Southwest. These two variables do not work well in either situation, yet they both show strong effects and significance levels at the nationwide level in Model 3. This is an important finding, though I am uncertain of the contributory factors.

The third variable hypothesized to have a negative effect on the TFR is percentage divorced. Its coefficient shows the anticipated effect and is statistically significant at $p < 0.005$. Percentage divorced was the weakest predictor among the five independent variables in all three nationwide fertility models. Nevertheless, it is one of the most powerful predictors in the three Northeast models and three Southwest models. Undoubtedly, this was one of the most unexpected discoveries in the entire analysis!

In Model 12, percentage minority has a positive relationship with fertility as expected and its coefficient is significant at $p < 0.005$. This signifies that for every one unit increase in the percentage of the population that is a member of one of China's minority nationalities, there is a corresponding increase in the TFR, holding the effects on the TFR of the other independent variables constant. This variable also had positive and significant associations in Models 3 and 9.

Percentage female illiterate has a positive and significant effect on the TFR as well at $p < 0.005$. The coefficients of percentage female illiterate in Models 3 and 9 were positive and had the same level of significance as the coefficient in Model 12. Thus, in addition to divorce, percentage minority and percentage female illiterate have also proven to be solid predictors of fertility among the 415 counties of the Southwest Region.

In my appraisal of the standardized coefficients, I see that percentage female illiterate has the strongest influence over the TFR, $b^* = 0.439$, followed sequentially by percentage minority, 0.350, and percentage divorced, -0.344. Together, the independent variables in Model 12 account for 64% of the variation in the TFR. The R^2 (adjusted) statistic is significant as F = 146.58 and p = 0.000.

I re-estimated Models 10, 11, and 12 with robust regression and found the robust results to be very similar to the OLS results. Thus, I am confident in my OLS findings. Tolerance values on each variable never dropped below 0.30; consequently, I do not believe multicollinearity threatened the accuracy of any of these models.

While percentage nonagriculture maintained significance, the other two economic development variables, percentage urban and percentage white collar, did not fare well in predicting fertility among the counties of the Southwest Region. The other variable which proved to be invalid was percentage one generation; the same variable which showed insignificance in all three of the Northeast models.

As I mentioned earlier, the effects of the economic development and percentage one generation variables may have been reduced in magnitude in Models 7-12 because of less variation in these variables among the two regional groups of counties. To check out this hypothesis, I compared the coefficient of relative variation (CRV) for the requisite variables in the total number of counties and the two regional groups of counties. The CRV is calculated by dividing the standard deviation by the mean. Indeed, my prediction that there would be more relative variation in the 2,432 counties than in the 216 Northeast counties and 415 Southwest counties proved to be correct. I found that the CRVs among the total group of counties in the variables were as follows: percentage nonagriculture, 0.247, percentage urban, 0.362, percentage white collar, 0.271, and percentage one generation, 0.324. These CRVs are higher than the corresponding CRVs among the Northeast counties of 0.157, 0.273, 0.212, 0.198, and among the Southwest counties of 0.229, 0.345, 0.257, and 0.289. Although several of the CRVs were comparable among these three groups, the greater relative variation among the variables in the total group of counties undoubtedly contributed to their superiority in predicting fertility in Models 1-6.

Three variables established themselves as especially robust predictors of fertility in the Southwest Region: percentage minority, percentage divorced, and percentage female illiterate. These three variables were also the most resilient variables in the fertility analysis of the counties in the Northeast. The fact that the weakest and strongest DTT variables are the same for both the high and low fertility regions is proof that our inclinations are not always correct.

These OLS results provide somewhat conflicting evidence regarding the efficiency of DTT in predicting Chinese fertility. I will discuss the conclusions and implications of the research conducted and reported in this thesis in the next chapter.

CHAPTER VII

CONCLUSIONS AND DISCUSSION

The main purpose of this thesis was to evaluate the levels of association between social and economic development variables and fertility rates among the counties of China in 2000. This is a noteworthy concern, given the sizeable fluctuations in fertility across China today. Earlier studies have established numerous links between various aspects of socioeconomic development and fertility decline in China (Birdsall and Jamison 1983; Tien 1984; Poston and Gu 1987; Freedman et al. 1988; Peng 1989; Poston and Jia 1990; Poston 2000). The results of my study are in keeping with these previous investigations, lending support to demographic transition theory, which states that lowered fertility and lowered mortality arise out of socioeconomic development. In this final chapter, I will summarize my research and findings, acknowledge what I could have done differently, and discuss needed future research in this area.

In order to sustain itself into the future, the People's Republic of China undertook in the 1970s a legendary demographic endeavor in the artificial constraint of population growth. Quite literally, the "later, longer, fewer" policy and the more stringent one-child policy were man-made efforts to speed up the demographic transition of the country. Their ultimate goal has been the stabilization and eventual decline of the population, via fertility at below-replacement levels for an extended period of time. This strategy has been triumphant in many ways because of the government's firm control over all aspects of Chinese life.

The "wan (later), xi (longer), shao (fewer)" policy, launched in 1971, insisted that couples marry late, extend their birth intervals, and have a small number of children. Urban couples were expected to delay marriage until age 25 for women and 28 for men and to have no more than two children (Attane 2002). Rural couples had a little more liberty, with marriage age minimums of 23 and 25, and a maximum of three children. This campaign was successful in reducing the birthrate from 40.0 per 1,000 in the 1960s to 18.3 per 1,000 in 1978 (Tien 1980).

In response to the nation's deteriorating economy, the post-Mao leadership began an extensive economic reform effort in 1979 called the "Four Modernizations" (Prybyla 1990). One of the most momentous outcomes of this project was the introduction of the "household responsibility system," which transformed communal farming into individual family farming (Feder et al. 1992; Prybyla 1990). Even though land parcels are extremely small, supervision is ubiquitous, and governmental contracts are obligatory, this new system granted peasants some small degree of autonomy.

Projections recommended that China restrict its population to 650-700 million by the middle of the twenty-first century to be economically and nutritionally stable. The government feared that the nation's young age cohorts would be an impediment to this goal and in January 1979, the government delivered the famous one-child policy (Banister 1984; Bongaarts and Greenhalgh 1985). This policy was intended limit the total population to 1.2 billion by the end of the 20^{th} century, so that the forthcoming population targets would be met.

Four factors were especially important in building civilian conformity to the one-child policy: (1) the "one child" certificate guaranteeing benefits to couples with one child who pledged to have no more, (2) the 1980 marriage law making birth control mandatory, (3) the escalating use of abortion and sterilization as fall back measures to accomplish the population planning targets, and (4) fines against families having third and higher parity births (Cooney and Li 1994; Banister 1987; Tien 1991; Bongaarts and Greenhalgh 1985). The policy seemed very promising to officials, who watched the TFR drop from 2.7 in 1981 to 2.1 in 1984 (Hardee-Cleaveland and Banister 1988).

The one-child policy has been extremely effective in urban areas. Not only are urban residents easier to monitor and control, small families are more conducive to their way of life. Quite the reverse, family planning in rural areas has been a challenge for officials, as demonstrated by the lofty percentages of unauthorized rural births. One reason for the struggle has been the "household responsibility system," which gave rural residents new financial independence from higher authorities. Moreover, by transferring responsibility from the commune to the household, the system increased the motivation for large families.

In an effort to strengthen family planning in rural areas, the State Family Planning Commission went through a series of leadership adjustments and developed new regulations regarding the one-child policy. At some points in this process, the Commission called for more leniency among the birth planning officials (i.e. cadres), while at others, they demanded tighter control. The result was considerable provincial variation in second-child permits, which has in turn contributed to the geographical variation in fertility seen in China today. According to Census 2000, the TFR ranged from a low of 0.67 in Beijing to a high of 2.19 in Guizhou Province.

Undoubtedly, China's family planning policy is largely responsible for the nation's current low fertility level of 1.22, as well as the spatial differences in fertility I just mentioned. Research has shown, however, that other factors have played a part in this fertility transition and subsequent variation at the regional, provincial, and county levels. In keeping with the expectations of demographic transition theory (DTT), quantitative studies conducted over the last twenty years have linked an assortment of socioeconomic factors with China's fertility decline and nationwide inconsistencies (Birdsall and Jamison 1983; Tien 1984; Poston and Gu 1987; Freedman et al. 1988; Peng 1989; Poston and Jia 1990; Poston 2000). My thesis was essentially a continuation of this work, using the new demographic data provided by Census 2000. I tested the efficiency of DTT variables in explaining the variation in the TFR among the counties of China. Specifically, I examined the ways in which variables such as ethnicity, agricultural detachment, urbanization, economic conditions, cultural norms and gender differences were related to Chinese fertility.

Though similar analyses have been done in the past, my thesis had some noteworthy features. The biggest advantage was that the data came from Census 2000, commended for its addition of short-form and long-form questionnaires. This meant that I was able to use predictors seldom used in previous investigations of Chinese fertility. These predictors were the percentage (15+) divorced and the percentage of one generation family households. Both variables became valuable components in different phases of my analysis.

An additional strength of this thesis is that I had previously performed an analysis of the variation in the TFR among the counties of Sichuan Province. This permitted me to identify and solve some methodological problems upfront, acting as a catalyst for my thesis.

According to Census 2000, the mean value of the TFR among the counties of China is 1.32 (SD .47), varying from a low of 0.41 in the Xiangyang district of Jiamusi city, Heilongjiang, to a high of 5.47 in Baqing County, Tibet. These wide ranging scores yield a range of 5.06, much wider than the range of the TFRs at the provincial level of 1.52. The four counties with the highest TFRs are located in the Tibet (Xizang) Autonomous Region, a poor, rural region with a large minority composition located in northwest China. The county with the lowest TFR is located in the center of Heilongjiang, a well developed province in northeast China. The sites of these outermost values seem suitable given the expectations of demographic transition theory (DTT). Please refer to **TABLE 1** for descriptive statistics of the independent variables. I will now provide an overview of my statistical analysis and discuss the relevant findings.

In this thesis, I estimated several Ordinary Least Squares regression equations to assess the effectiveness of demographic transition theory in predicting Chinese fertility. Before the OLS analysis, I employed many diagnostics to ensure statistical accuracy in my models. An evaluation of the distributions of the dependent and independent variables led me to conduct power transformations on all but one of the variables to bring in the positive outliers. Most variables were transformed by the base 10 logarithm. I also assessed the zero-order correlations of the independent variables with the TFR and with each other (see **TABLE 2**). Scatterplots of these bivariate relationships were then examined, which provided a visual representation of the positive and negative associations among the variables (see **FIGURES 1-10**). These initial tests showed some collinearity in my independent variables. Thus I decided to remove percentage illiteracy, percentage male illiteracy, and female status from the group of independent variables. This resulted in seven predictors to be used throughout my OLS analysis: (1) percentage urban, (2) percentage nonagriculture, (3) percentage white collar, (4) percentage one generation, (5) percentage divorced, (6) percentage minority, and (7) percentage female illiterate.

In consideration of the three highly correlated economic development variables, I elected to divide my analysis into three separate OLS models (see **TABLE 3**). Models 1-3 included five independent variables and the TFR. The four cultural, minority, and illiteracy independent variables were exactly the same in all three models; the single economic variable was the distinguishing feature.

Models 1-3 examined the impact of these independent variables on the TFR among 2,432 counties across China. The coefficients of the independent variables in Model 1 were all significant and showed the hypothesized effect on the TFR. With the exception of percentage divorced, the coefficients in Model 2 were also significant and signed in the expected direction. And finally, the coefficients in Model 3 revealed that all of the independent variables have significant effects on fertility in the hypothesized direction.

My results from Models 1 and 2 indicate that the classic DTT variables, percentage urban and percentage nonagriculture, have the strongest relative effects on fertility. The success of these particular variables champions the notions of demographic transition theory. My findings from Model 3 showed that percentage female illiterate has more influence over fertility than the other independent variables, percentage white collar, percentage one generation, percentage minority, and percentage divorced. Because there are so few white collar workers among the counties of China, this economic development variable exerted less comparative influence than those in the first two models.

The DTT variables that were used in Models 1-3 were identical in Models 4-6, the only difference being the addition in Models 4-6 of the region dummies (see **TABLE 4**). DTT performed well even under the regional controls, as the DTT coefficients in Models 4-6 were very comparable to those in Models 1-3. With the exclusion of percentage divorced, the coefficients of the DTT variables showed the hypothesized effects on the TFR and were statistically significant in all three models. Additionally, the coefficients of the region variables were all positive and significant. This means that even after taking into account the DTT variables in each of the OLS equations, there continue to be differences in fertility among the regions.

In view of these findings, I chose to conduct two region-specific fertility analyses, one on the highest fertility region of China and one on the lowest (see

TABLES 5-6). This was intended to test the predictive power of the DTT variables in two divergent settings, each being composed of a rather homogeneous group of low (or high) fertility counties. The impact of the DTT variables on fertility changed drastically under these restricted conditions. Models 7-12 are analogous to the basic three models seen in **TABLES 1-3**.

TABLES 5-6 display the results regarding the fertility analyses of the Northeast Region and Southwest Region. I will begin with the two predictors which floundered in Models 7-12. The coefficient of percentage one generation households shows neither the hypothesized relationship nor a statistically significant effect on the TFR in any of the six models. This variable had a significant negative effect on fertility in all three nationwide models, but when operating on fertility only in the counties within each of these regions, its effect is reversed and becomes insignificant.

Percentage white collar was included in Models 9 and 12, and although it has the expected negative relationship with the TFR, its coefficient is insignificant in both models. When this variable was included in Model 3, it had one of the strongest relative effects among the five predictors. I speculate that this variable may have succeeded in the nationwide analysis simply because of its association with urbanization and economic progress. In the nationwide analysis, there was sufficient developmental variation among the 2,432 counties for this variable to perform well in the OLS regression equation. Variation was smaller among the 415 counties in the underdeveloped Southwest and also amid the 216 counties in the developed Northeast. The predictive capabilities of percentage white collar were thus effectively removed.

The coefficients of percentage nonagriculture, percentage minority, percentage divorced, and percentage female illiterate show the hypothesized effects and are statistically significant in Models 7-9. The coefficient of percentage urban is insignificant in Model 10 of the Southwest Region, but shows a negative and significant effect on fertility in Model 7 of the Northeast Region. Percentage divorced has the largest relative impact on fertility in the developed Northeast in two out of the three models. This is a fascinating turn of events given that the variable was the weakest predictor in all three nationwide models. In all three of the Southwest models, it is percentage female illiterate that has the largest relative effect on fertility. Moreover, each of these variables, percentage divorced and percentage female illiterate, perform exceptionally well in every one of the high and low fertility models. Percentage minority should also be mentioned for maintaining significance and strength of effect in all six region-specific models.

Even though the effects of the DTT predictors changed somewhat from the nationwide to the region-specific fertility analyses, overall I think they performed quite nicely. The OLS results from the twelve different models convey that demographic transition theory is applicable for predicting and understanding fertility among the counties of China. Irrespective of the nation's extensive family planning policy, it is beyond doubt that other factors contribute to the inconsistent fertility rates across the country.

Looking back, there are faults in my analysis which I will now take the time to discuss. I will also offer suggestions for how these issues could have been more effectively managed.

My analysis could have been greatly enhanced had it been possible to include a measurement of per capita income. Previous Chinese fertility studies have benefited from the use of such indicators, since they provide a more truthful account of economic status. The three economic development variables I used, percentage urban, percentage nonagriculture, and percentage white collar, are related to economic standing, but an indicator of tangible economic standing would have been better. Census 2000 does not contain such information, so I am restricted in this way by the data.

The analysis also would have benefited from the incorporation of a few family planning variables, given that the family planning program has such a huge effect on fertility in China, apart from and in conjunction with socioeconomic factors. For example, an independent variable pertaining to family planning costs might have been particularly useful, since wealthier jurisdictions often have the upper hand when it comes to birth restriction. Their clinics are better outfitted, but more importantly they are able to offer greater financial incentives to couples accepting the one-child certificate. Moreover, it would have been fascinating to bring in a measurement of abortion frequency, since abortions are routinely used to accomplish the birth planning targets. Census 2000 supplies no information regarding family planning activities; so again, I am constrained in this respect.

Another variable that should have been included in this analysis is the infant mortality rate (IMR), which has been found to have a large impact on fertility in China. A major postulate of demographic transition theory (DTT) is that fertility decline follows infant mortality decline. I thus consider it to be a necessary component in any study which attempts to evaluate the relevancy of DTT in projecting fertility outcomes. Census 2000 does not provide mortality rates for each age group; however, I could have resorted to another measure capturing the survival rate of children. The census has data on "children ever born, per woman 15-50" and also on "the average number of children survived until the date of the census." While the measure would not have been a satisfactory substitute for the IMR, it is likely the closest approximation afforded by the data source, and it should have been employed. Future analysis would do well to use such a proxy.

The one problematic variable in this analysis was the female status variable, which is a ratio of percentage female illiterate to percentage male illiterate. A value greater than 1.0 on this variable signified more female illiteracy, while a value less than 1.0 signified more male illiteracy. Educational inequality encourages social inequality, which demographers have linked to higher levels of fertility. Therefore, I hypothesized that this variable would be positively associated with Chinese fertility. Both its correlation with the TFR and its effect on the TFR were negative, which contradicted my expectations and the claims of DTT. Obviously, the female status variable failed to test the gender bias-fertility association in this county-level analysis of China. Perhaps I should have resorted to a different type of measurement in my attempts to quantify sexual inequality. The census provides data on educational attainment for both men and women, with categories ranging from “never been to school” to “graduate school.” I could have created an ordinal variable consisting of the eight categories of educational attainment provided by the census, from 1 = lowest, to 8 = highest. I have reservations about such a variable; however, since it is debatable as to whether or not the educational categories are evenly spaced. For instance, it is questionable whether a transition from high school to technical school is equivalent to a transition from college to graduate school. Data are also available on “the average years of education” for men and women. So another alternative would have been to develop a female status variable from the ratio of these two averages. One more option would have been to create a series of education dummy variables.

A further criticism that could be put forth about my analysis is the absence of indexes, which may have helped with variable augmentation and improvement. For example, in this thesis I estimated the effects of percentage of the population in nonagriculture, percentage of the population urban, and percentage of the labor force in white collar jobs in separate OLS models because of the high degree of collinearity that exists among them. Instead, it may have been better to combine them into a single index. Uniting them in this way would bypass collinearity issues and allow for an estimate of their simultaneous effect on fertility.

In the years to come, I hope to extend my research on Chinese fertility by utilizing more involved statistical strategies. In the present analysis, I predicted the average fertility behavior of women in counties with the contextual characteristics of the counties. This is a common tactic and it has proven to be a reasonably successful one. There are many people who strongly favor this macro-approach to the study of fertility. This group contends that it is the larger environment that shapes fertility decisions. In other words, it is the contextual characteristics of the community in which a person lives that ultimately determine the number of children ever born to him/her. For this reason, they believe it makes more sense to measure the fertility outcomes of the aggregate, rather than those of the individual. The variables used in my thesis, such as percentage urban and percentage nonagriculture, are good examples of these macro-level predictors. In his article, entitled "Where Do Babies Come From?" Norman B. Ryder expressed this macro position regarding fertility research (1980: 197): "Fertility is a collective property of a system, to be explained in relation to other collective properties of the same system. Far from being an expedient in lieu of individual-level analysis, a macro orientation is theoretically appropriate" (1980: 197).

At the same time, others argue that since fertility decisions are made at the individual-level, they should be measured and predicted at the individual-level, as opposed to the aggregate-level. In other words, it is individuals that generate the fertility levels of the populace, rather than the other way around. Therefore, in an attempt to understand that which influences fertility, one should look to the individual for answers. A person's educational attainment, ethnic affiliation, and occupational status are examples of predictors that may be used in this type of micro-level study.

Each perspective is valid and worthy of pause. Rather than disregarding either line of reasoning altogether, there is an approach that satisfies both factions of logic—

the multilevel analysis. In a multilevel analysis, there are independent variables at more than one level predicting the outcomes of the dependent variable. This allows the model to take into account both the individual-level (i.e. micro) characteristics as well as the aggregate-level (i.e. macro) characteristics in producing estimates.

In future research, I aspire to use multilevel modeling in my dissertation on Chinese fertility, as it will impart a more comprehensive understanding of the dynamics involved in fertility decisions and outcomes. The 1% Sample of the 2000 Census of China will soon be released, which provides descriptive data on a sample of women from all the counties in China. Micro-level variables that may be used regarding the individual woman are age, agricultural affiliation (yes/no), level of education, and whether she is a member of the Han majority nationality (yes/no). Macro-level variables that may be used regarding the county context include the percentage of persons employed in non-agricultural jobs, the proportion of persons in each county working in white collar jobs, and the percentage of persons (15+) illiterate. The outcome variable will likely be a dummy variable indicating whether the woman had a birth during the 18 months preceding the date of the census of November 1, 2000, 1=Yes, 0=No. I believe my fertility research will benefit from this more progressive statistical strategy, and in turn, I will have more to contribute to the field.

Demographers have expended a great deal of time and energy in explaining that which they once thought was virtually impossible, the manifestation of below-replacement fertility and its rapid movement through societies across the world. But today, there is an epidemic of even greater demographic concern that leaves our economies and social well-being in the balance. This is the spread of lowest-low fertility, or TFR levels of 1.3 or less. Fourteen countries in Southern, Central, and Eastern Europe arrived at this lowest-low fertility level in the 1990s, the initial qualifiers being Spain and Italy (Kohler, Billari, and Ortega 2002). As Kohler et al. (2002: 642) explain, "If the TFR declines further and persists at a level of 1.0, the annual rate of decline in the stable population rises to 2.4 percent and the halving times of population size and birth cohorts are merely 29 years."

While these European countries have TFRs of 1.3, most African countries today boast TFRs of 6.0. The demographic research regarding these high fertility societies is limited because most sub-Saharan African countries did not perform an exhaustive census until around 1960. Data are still very restricted in comparison to other countries around the world, but there have been many improvements over the last three decades. Purportedly, all African countries have TFRs above 4.0, and only a handful have TFRs below 6.0 (Foote et al. 1993).

There is an incredible paradox in the world's fertility. If we are to ever gain a true understanding this most complicated demographic phenomenon, it will be by way of cross-country comparisons. China is fascinating in that within itself it exhibits high fertility levels resembling those of sub-Saharan Africa, as well as lowest-low fertility levels resembling those of Europe. Perhaps some of the same socioeconomic and cultural factors inspiring these divergent rates of fertility are shared among China, Africa, and Europe. It is imperative that demographers address this question in the years to come and I hope to be involved in the challenge.

REFERENCES

Attane, Isabelle. 2001. "Chinese Fertility on the Eve of the 21st Century: Fact and Uncertainty." *Population: An English Selection* 13:71-100.

Attane, Isabelle. 2002. "China's Family Planning Policy: An Overview of Its Past and Future." *Studies in Family Planning* 33:103-113.

Banister, Judith. 1984. "Population Policy and Trends in China, 1978-83." *The China Quarterly* 100:717-741.

Banister, Judith. 1987. *China's Changing Population*. Stanford: Stanford University Press.

Beauty: The Land of China. (1998). Liaoning. Retrieved January 2005, from the ThinkQuest Web site: http://library.thinkquest.org/20443/liaoning.html

Beijing Domestic Service. 1985. "Wang Wei on Publicizing Family Planning in Beijing," January 11, 1985, in FBIS, *Daily Report—China*, no. 010 (January 15, 1985): K27.

Beijing Review. 1984. "Education Level of Population," April 27, 1984, in *Beijing Review* 27:22.

Birdsall, Nancy and Dean T. Jamison. 1983. "Income and Other Factors Influencing Fertility in China." *Population and Development Review* 9:651-675.

Bongaarts, John and Susan Greenhalgh. 1985. "An Alternative to the One-Child Policy in China." *Population and Development Review* 11:585-617.

Caldwell, John C. 1976. "Toward a Restatement of Demographic Transition Theory." *Population and Development Review* 2:321-366.

Caldwell, John C. 1979. "Age as a Factor in Mortality Decline: An Examination of Nigerian Data." *Population Studies* 33:395-413.

Chan, Kam Wing. 2003. "Chinese Census 2000: New Opportunities and Challenges." *The China Review* 3:1-12.

Chen, Muhua. 1979. "For the Realization of the Four Modernizations, There Must be Planned Control of Population Growth," *Renmin Ribao (People's Daily)*, in *Population and Development Review* 5:723-730.

China Economic Information Network (CEInet), P.R. China. Provinces of the P.R.C. (n.d.). Retrieved January 2005, from http://ce.cei.gov.cn/region/

ChinaPlanner.com. Wild West of China: Linxia and Around. (n.d.). Retrieved January 2005, from http://www.chinaplanner.com/westchina/gansu/linxia.htm

China Population Information and Research Center (CPIRC). (n.d.). Retrieved January 2005, from http://www.cpirc.org.cn/en/welcome.htm

Chinastage.com. Ningxia. (1998). Retrieved January 2005, from http://www.chinastage.com/travel-guide/Ningxia.html

ChinaTravel.com. China Destination Guides. (n.d.). Retrieved January 2005, from http://www.chinatravel.com/index.htm

Coale, Ansley J. 1973. "The Demographic Transition Reconsidered." Pp. 53-72 in *Proceedings: International Population Conference, Liege, 1973*. Liege, Belgium: International Union for the Scientific Study of Population.

Cooney, Rosemary Santana and Jiali Li. 1994. "Household Registration Type and

Compliance with the 'One Child' Policy in China, 1979-1988." *Demography* 31: 21-32.

Daily Reports: The Comprehensive Source. (2001, March). Chinese Census Shows Growth Rate Decreasing, Still More Boys Than Girls Born Overall. Retrieved January 2005, from the Henry Kaiser Family Foundation Web site: http://www.kaisernetwork.org/Daily_reports/rep_index.cfm?DR_ID=3704

Davis, Kingsley. 1963. "The Theory of Change and Response in Modern Demographic History." *Population Index* 29:345-66.

Feder, Gershon, Lawrence J. Lau, Justin Y. Lin, and Xiaopeng Luo. 1992. "The Determinants of Farm Investment and Residential Construction in Post-Reform China." *Economic Development and Cultural Change* 41:1-26.

Foote, Karen A., Kenneth H. Hill, and Linda G. Martin (Eds.). 1993. *Demographic Change in Sub-Saharan Africa*. Washington, DC: The National Academies Press (NAP). Retrieved February 2005, from the NAP website: http://www.nap.edu/

Freedman, Ronald, Xiao Zhenyu, Li Bohua, and William Lavely. 1988. "Local Area Variations in Reproductive Behavior in the People's Republic of China, 1973-1982." *Population Studies* 42:39-57.

GreatestCities.com. Anhui Province, Bengbu, pictures. (2001). Retrieved January 2005, from http://www.greatestcities.com/Asia/China/Anhui_Province/

Greenhalgh, Susan. 1986. "Shifts in China's Population Policy." *Population and Development Review* 12:491-515.

Hamilton, Lawrence C. 1992. *Regression with Graphics: A Second Course in Applied*

Statistics. Belmont, CA: Duxbury Press.

Hardee-Cleaveland, Karen and Judith Banister. 1988. "Fertility Policy and Implementation in China, 1986-88." *Population and Development Review* 14: 245-286.

Hirschman, Charles. 1994. "Why Fertility Changes." *Annual Review of Sociology* 20: 203-233.

Information Office of the State Council of the People's Republic of China. *Regional Ethnic Autonomy in Tibet*. (May 24, 2004). Retrieved on January 2005, from http://servizi.radicalparty.org/documents/index.php?func=detail&par=3513

Invest-inchina.com. Brief Introduction and Investment Environment of Hebei (November 27, 2002). Retrieved January 2005, from http://www.invest-inchina.com/new/en/provnces%20and%20cities/page_84.htm

Italtrade.com: The Official Gateway to Italian Business and Trade. Ningxia. (n.d.). Retrieved January 2005, from http://www.ice.gov.it/estero2/pechino/profningxia

Johnson, D. Gale. 2000. "Agricultural Adjustment in China: Problems and Prospects." *Population and Development Review* 26:319-334.

Kaplan, Hillard. 1994. "Evolutionary and Wealth Flows Theories of Fertility." *Population and Development Review* 20:753-791.

Kelly, William R., Dudley L. Poston, Jr., and Phillips Cutright. 1983. "Determinants of Fertility Levels and Change among Developed Countries: 1958-1978." *Social Science Research* 12:87-108.

Kennedy, Bingham, Jr. 2001. "Dissecting China's 2000 Census." *Population Today*

June:1-3.

Knodel, John and Etienne van de Walle. 1979. "Lessons from the Past: Policy Implications of Historical Fertility Studies." *Population and Development Review* 5:217-245.

Kohler, Hans-Peter, Francesco C. Billari, and Jose Antonio Ortega. 2002. "The Emergence of Lowest-Low Fertility in Europe during the 1990s." *Population and Development Review* 28:641-680.

Landsberger, Stefan R. Fifth National Census. (n.d.). Retrieved January 2005, from Stefan Landsberger's Poster Pages: http://www.iisg.nl/~landsberger/rkpc.html

Lardy, Nicholas R. 1983. *Agriculture in China's Modern Economic Development*. Cambridge: Cambridge University Press.

Lavely, William. 2001. "First Impressions from the 2000 Census of China." *Population and Development Review* 27:755-769.

Liu, Mary. (2004, May). Dongze Joy Foreign Language School. Retrieved January 2005, from http://maryliu.fotopages.com/?entry=98658

Liu, Zheng and Jian Song. 1981. *China's Population: Problems and Prospects*. Beijing, China: New World Press.

London, Bruce and Kenneth Hadden. 1989. "The Spread of Education and Fertility Declines: A Thai Province Level Test of Calwell's 'Wealth Flows' Theory." *Rural Sociology* 54:17-36.

Mason, Karen Oppenheim. 1997. "Explaining Fertility Transitions." *Demography* 34:443-454.

Mauldin, W. Parker, Bernard Berelson, and Zenas Sykes. 1978. "Conditions of Fertility Decline in Developing Countries, 1965-75." *Studies in Family Planning* 9:89-147.

Noll, Paul. China Provinces, Autonomous Regions and Municipalities. (1997). Retrieved January 2005, from http://www.paulnoll.com/China/Provinces/ index.html

Notestein, Frank. 1953. "Economic Problems of Population Change." Pp. 13-31 in *Proceedings of the Eighth International Conference of Agricultural Economists*. London: Oxford University Press.

Peiyun, Peng. 1988. "To Eradicate the Undesirable Marriage Customs and to Promote the Construction of Socialist Spirit and Civilization." (in Chinese) *China Population* 29:1.

Peng, Xizhe. 1989. "Major Determinants of China's Fertility Transition." *The China Quarterly* 117:1-37.

Poston, Dudley L., Jr. 2000. "Social and Economic Development and the Fertility Transitions in Mainland China and Taiwan." *Population and Development Review* 26:40-60.

Poston, Dudley L., Jr. and Baochang Gu. 1987. "Socioeconomic Development, Family Planning, and Fertility in China." *Demography* 24:531-551.

Poston, Dudley L., Jr. and Zhongke Jia. 1990. "Socioeconomic Structure and Fertility in China: A County Level Investigation." *Journal of Biosocial Science* 22:507-515.

Prybyla, Jan S. 1990. "Economic Reform of Socialism: The Dengist Course in China." *Annals of the American Academy of Political and Social Science* 507:113-122.

Qian, Zhenchao. 1997. "Progression to Second Birth in China: A Study of Four Rural Counties." *Population Studies* 51:221-228.

Riskin, Carl. 1987. *China's Political Economy: The Quest for Development since 1949*. New York: Oxford University Press.

Ryder, Norman B. 1980. "Where Do Babies Come From?" Pp. 189-202 in *Sociological Theory and Research: A Critical Appraisal*, edited by H.M. Blalock. New York: The Free Press.

Scharping, Thomas. 2003. *Birth Control in China 1949-2000: Population Policy and Demographic Development*. New York: RoutledgeCurzon.

Shuimin, Li and Beiliu People's Government. Internet Dialogue on Ecological Sanitation. (November 15-December 20, 2001). *Accelerate the Development of Rural Dry Ecological Sanitation*. Retrieved January 2005, from http://www.ias.unu.edu/proceedings/icibs/ecosan/li.html

Sobotka, Tomas. 2004. "Is Lowest-Low Fertility in Europe Explained by the Postponement of Childbearing?" Population and Development Review 30:195-220.

State Council and State Statistical Bureau. 2002. Population Census Office Under the State Council and Department of Population Statistics, State Statistical Bureau. *Tabulation on the 2000 Population Census of the People's Republic of China*. Beijing, China: China Statistical Publishing House.

Thompson, Warren S. 1929. “Population.” *The American Journal of Sociology* 34:959-975.

Tien, H. Yuan. 1980. “Wan, Xi, Shao: How China Meets Its Population Problem.” *International Family Planning Perspectives* 6:65-70.

Tien, H. Yuan. 1984. “Induced Fertility Transition: Impact of Population Planning and Socioeconomic Change in the People’s Republic of China.” *Population Studies* 38:385-400.

Tien, H. Yuan. 1991. *China’s Strategic Demographic Initiative*. New York: Praeger.

Verber, Mark. Xinjiang Uyghur Autonomous Region, China. (n.d.). Retrieved on January 2005, from http://www.veber.com/mark/travel/xinjiang

Walder, Andrew G. 1989. “Social Change in Post-Revolution China.” *Annual Review of Sociology* 15:405-424.

Weaver, Lisa. (2000, November). Millions Head Out to Count China’s Population. *CNN, The Associated Press, and Reuters*. Retrieved January 2005, from http://archives.cnn.com/2000/ASIANOW/east/11/02/china.census/

Wikipedia: The Free Encyclopedia. Guangdong. (n.d.). Retrieved January 2005, from http://en.wikipedia.org/wiki/Guangdong

Zeng, Yi, T. Paul Schultz, Wang Deming, and Gu Danan. (2002, August). Association of Divorce with Socio-Demographic Covariates in China, 1955-1985: Event History Analysis Based on Data Collected in Shanghai, Hebei, and Shaanxi. *Demographic Research* 7:1-28. Retrieved January 2005, from http://www.demographic-research.org/volumes/vol7/11/1-11.pdf

APPENDIX A

TABLES

TABLE 1 Descriptive Statistics for Fertility Rates and Social, Economic, and Cultural Variables: 2,870 Counties of China, 2000

Variable	Mean	Standard deviation	Minimum value	Maximum value
Total fertility rate	1.32	0.47	0.41 Xiangyang District of Jiamusi City (Heilongjiang)	5.47 Baqing County (Tibet)
% Population Nonagriculture	27.84	25.25	2.07 Jiangda County (Tibet)	98.91 Tiedong District of Anshan City (Liaoning)
% Population Urban	28.66	20.96	0.56 Linxia County (Gansu)	100.00 3 Districts of Bengbu City (Annui) Meijiang District of Meizhou City (Guangdong)
% Labor Force White Collar	8.59	6.49	1.37 Dongxiangzu Zizhi County (Gansu)	38.62 Xiangyang District of Jiamusi City (Heilongjiang)
% Population (15+) Divorced	1.09	0.79	0.16 Hongsibao District of Wuzhong City (Ningxia)	8.23 Leiwuqi County (Tibet)
% Households One Generation	20.93	7.05	3.75 Dongxiangzu Zizhi County (Gansu)	58.87 Baoan District of Shenzhen City (Guangdong)
% Population Minority	16.14	28.97	0.00 Qingjian, Zizhou, Lin and Shilou Counties of (Shanxi)	99.78 Angren County (Tibet)
% Population (15+) Illiterate	11.49	11.19	0.55 Beiliu City (Guangxi)	86.22 Baqing County (Tibet)
% Male Population (15+) Illiterate	6.76	8.55	0.24 Beiliu City (Guangxi)	78.16 Baqing County (Tibet)
% Female Population (15+) Illiterate	16.47	14.27	0.89 Beiliu City (Guangxi)	94.33 Baqing County (Tibet)
Female Status	3.16	1.41	0.94 Kangbao County (Hebei)	16.49 Pingyuan County (Guangdong)

SOURCE: 2000 Population Census of China (State Council and State Statistical Bureau 2002)

TABLE 2 Zero-order Correlations of Dependent and Independent Variables: 2,432 Counties and County Equivalents of China, 2000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) log_{10} Total Fertility Rate (TFR)	__										
(2) log_{10} % Population Non-agriculture	-0.568	__									
(3) sqrt % Population Urban	-0.560	0.845	__								
(4) log_{10} % Labor Force White Collar	-0.455	0.838	0.754	__							
(5) % Households One Generation	-0.422	0.464	0.538	0.449	__						
(6) log_{10} % Population (15+) Divorced	-0.087	0.328	0.228	0.363	0.178	__					
(7) log_{10} % Population Minority	0.232	0.035	-0.028	0.056	-0.102	0.440	__				
(8) log_{10} % Population (15+) Illiterate	0.439	-0.428	-0.386	-0.298	-0.179	0.113	0.298	__			
(9) log_{10} % Male Population (15+) Illiterate	0.394	-0.406	-0.399	-0.261	-0.190	0.206	0.314	0.941	__		
(10) log_{10} % Female Population (15+) Illiterate	0.441	-0.412	-0.361	-0.300	-0.167	0.069	0.284	0.987	0.878	__	
(11) log_{10} Female Status	-0.115	0.184	0.249	0.064	0.127	-0.313	-0.197	-0.375	-0.665	-0.228	__

SOURCE: 2000 Population Census of China (State Council and State Statistical Bureau 2002)

TABLE 3 Multiple Regression Coefficients for the TFR on DTT Independent Variables: 2,432 Counties and County Equivalents of China, 2000

Independent Variables	Model 1	Model 2	Model 3
Sqrt % Population Urban	-0.060*** (-0.371)		
Log_{10} % Population Nonagriculture		-0.175*** (-0.405)	
Log_{10} % Labor Force White Collar			-0.158*** (-0.259)
% Households One Generation	-0.007*** (-0.154)	-0.008*** (-0.181)	-0.011*** (-0.236)
Log_{10} % Population (15+) Divorced	-0.035*** (-0.065)	-0.008 (0.015)	-0.023* (-0.042)
Log_{10} % Population Minority	0.020*** (0.167)	0.021*** (0.179)	0.019*** (0.161)
Log_{10} % Female Population (15+) Illiterate	0.120*** (0.238)	0.098*** (0.194)	0.141*** (0.281)
Constant	0.382	0.651	0.380
R-squared (adjusted)	0.421	0.434	0.381

* P < 0.05, ** P < 0.01, *** P < 0.005

SOURCE: 2000 Population Census of China (State Council and State Statistical Bureau 2002)

TABLE 4 Multiple Regression Coefficients for the TFR on DTT Independent Variables and Region: 2,432 Counties and County Equivalents of China, 2000

Independent Variables	Model 4	Model 5	Model 6
Sqrt % Population Urban	-0.051*** (-0.315)		
Log_{10} % Population Nonagriculture		-0.139*** (-0.322)	
Log_{10} % Labor Force White Collar			-0.123*** (-0.201)
% Households One Generation	-0.008*** (-0.183)	-0.010*** (-0.209)	-0.011*** (-0.248)
Log_{10} % Population (15+) Divorced	0.005 (0.009)	0.018 (0.033)	0.007 (0.013)
Log_{10} % Population Minority	0.020*** (0.168)	0.020*** (0.171)	0.020*** (0.169)
Log_{10} % Female Population (15+) Illiterate	0.086*** (0.172)	0.076*** (0.152)	0.103*** (0.206)
North	0.244*** (0.299)	0.223*** (0.273)	0.295*** (0.361)
East	0.242*** (0.342)	0.207*** (0.292)	0.255*** (0.359)
Central & South	0.282*** (0.396)	0.253*** (0.355)	0.293*** (0.412)
Southwest	0.352*** (0.451)	0.323*** (0.415)	0.367*** (0.470)
Northwest	0.214*** (0.262)	0.194*** (0.238)	0.260*** 0.319
Constant	0.215	0.420	0.168
R-squared (adjusted)	0.495	0.494	0.461

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$

SOURCE: 2000 Population Census of China (State Council and State Statistical Bureau 2002)

TABLE 5 Multiple Regression Coefficients for the TFR on DTT Independent Variables: 216 Counties and County Equivalents of Northeast China, 2000

Independent Variables	Model 7	Model 8	Model 9
Sqrt % Population Urban	-0.024* (-0.226)		
Log_{10} % Population Nonagriculture		-0.081** (-0.242)	
Log_{10} % Labor Force White Collar			-0.044 (-0.106)
% Households One Generation	0.004 (0.094)	0.003 (0.074)	0.002 (0.054)
Log_{10} % Population (15+) Divorced	-0.165*** (-0.370)	-0.147*** (-0.330)	-0.198*** (-0.445)
Log_{10} % Population Minority	0.047*** (0.264)	0.044*** (0.251)	0.050*** (0.282)
Log_{10} % Female Population (15+) Illiterate	0.190*** (0.317)	0.198*** (0.330)	0.202*** (0.337)
Constant	-0.458	-0.324	-0.503
R-squared (adjusted)	0.449	0.450	0.437

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$

SOURCE: 2000 Population Census of China (State Council and State Statistical Bureau 2002)

TABLE 6 Multiple Regression Coefficients for the TFR on DTT Independent Variables: 415 Counties and County Equivalents of Southwest China, 2000

Independent Variables	Model 10	Model 11	Model 12
Sqrt % Population Urban	-0.006 (-0.029)		
Log_{10} % Population Nonagriculture		-0.083*** (-0.147)	
Log_{10} % Labor Force White Collar			-0.052 (-0.063)
% Households One Generation	0.001 (0.019)	0.004 (0.067)	0.002 (0.035)
Log_{10} % Population (15+) Divorced	-0.262*** (-0.363)	-0.230*** (-0.319)	-0.248*** (-0.344)
Log_{10} % Population Minority	0.041*** (0.335)	0.043*** (0.348)	0.043*** (0.350)
Log_{10} % Female Population (15+) Illiterate	0.245*** (0.437)	0.219*** (0.391)	0.247*** (0.439)
Constant	-0.343	-0.142	-0.317
R-squared (adjusted)	0.636	0.643	0.637

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$

SOURCE: 2000 Population Census of China (State Council and State Statistical Bureau 2002)

APPENDIX B

FIGURES

FIGURES 1-10 Scatterplots of TFR and Independent Variables: 2,432 Counties and County Equivalents of China, 2000

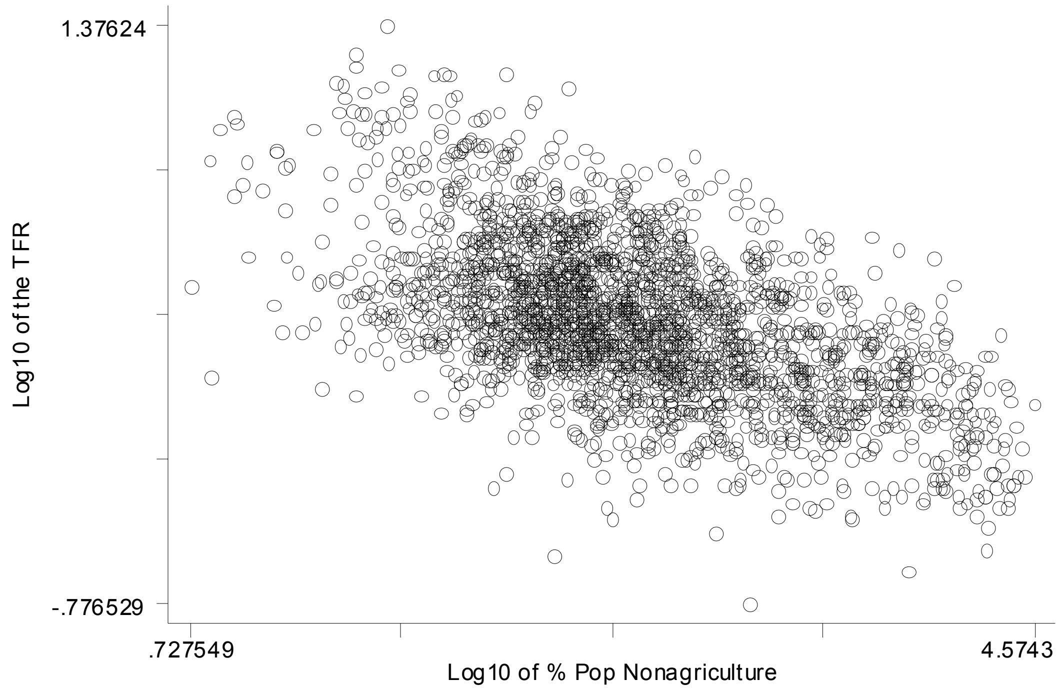

FIGURE 1: TFR & Percentage Population Nonagriculture (r = -0.568)

FIGURE 2: TFR & Percentage Population Urban (r = -0.560)

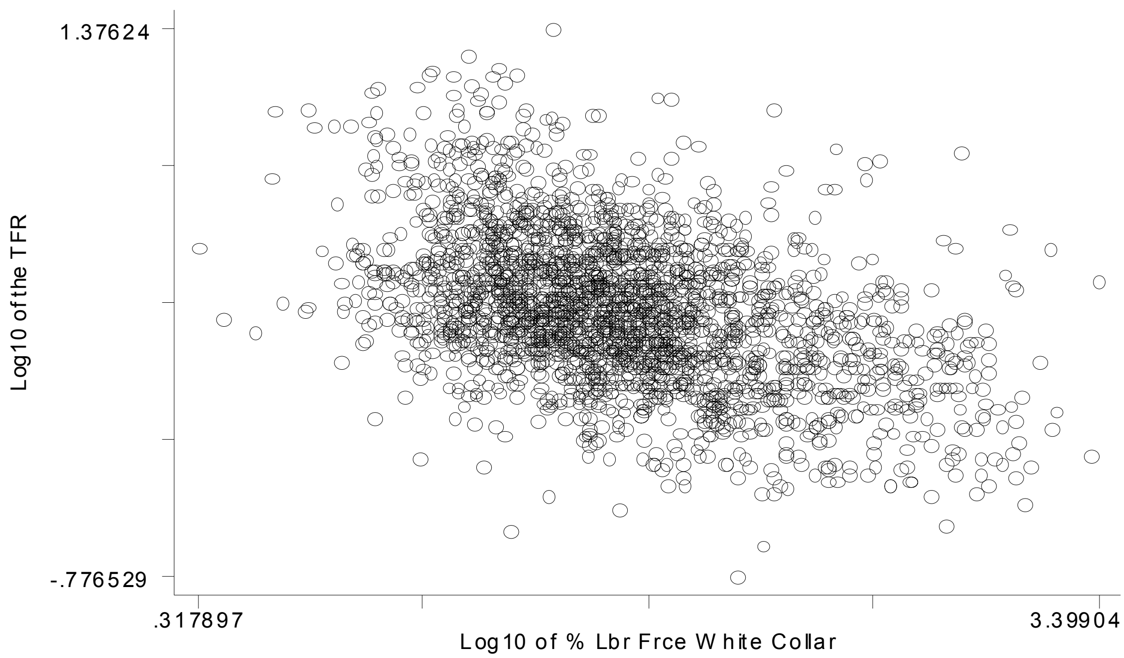

FIGURE 3: TFR & Percentage Labor Force White Collar (r = -0.455)

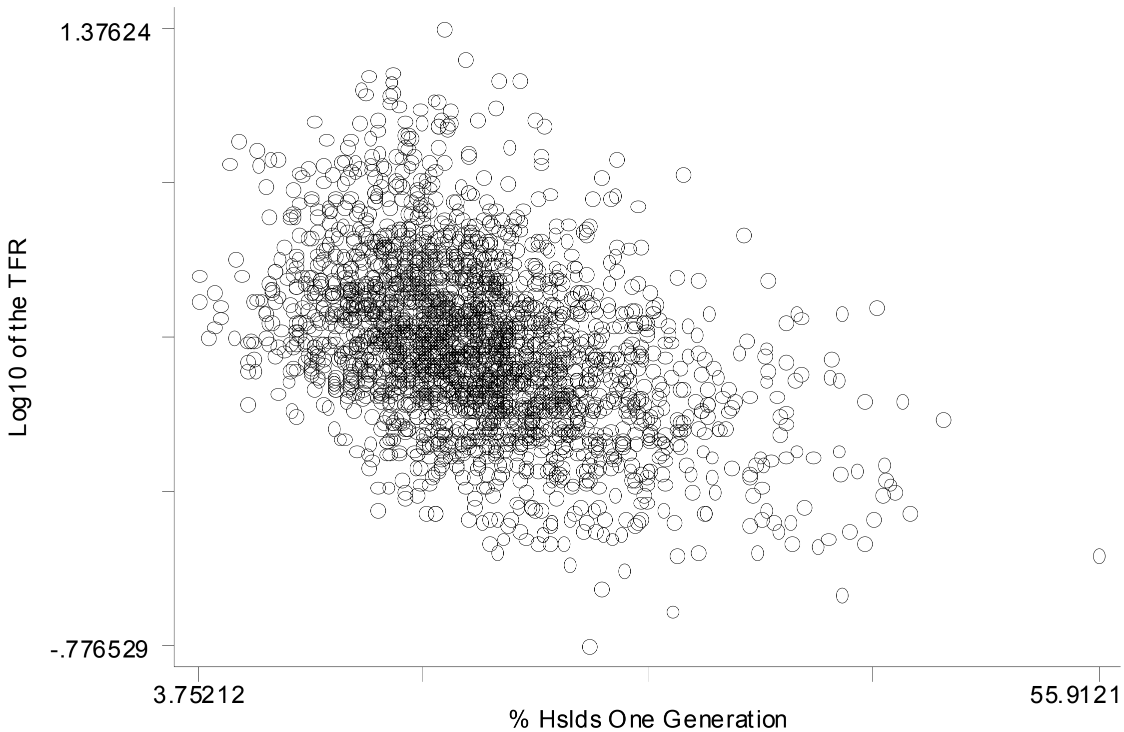

FIGURE 4: TFR & Percentage Households One Generation (r = -0.422)

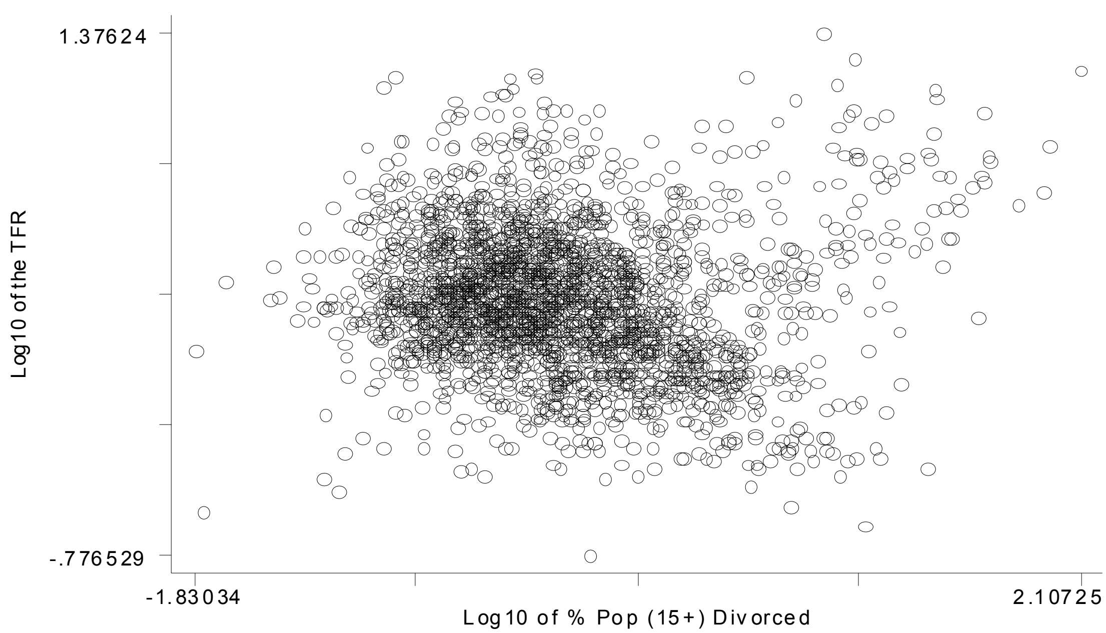

FIGURE 5: TFR & Percentage Population (15+) Divorced (r = -0.087)

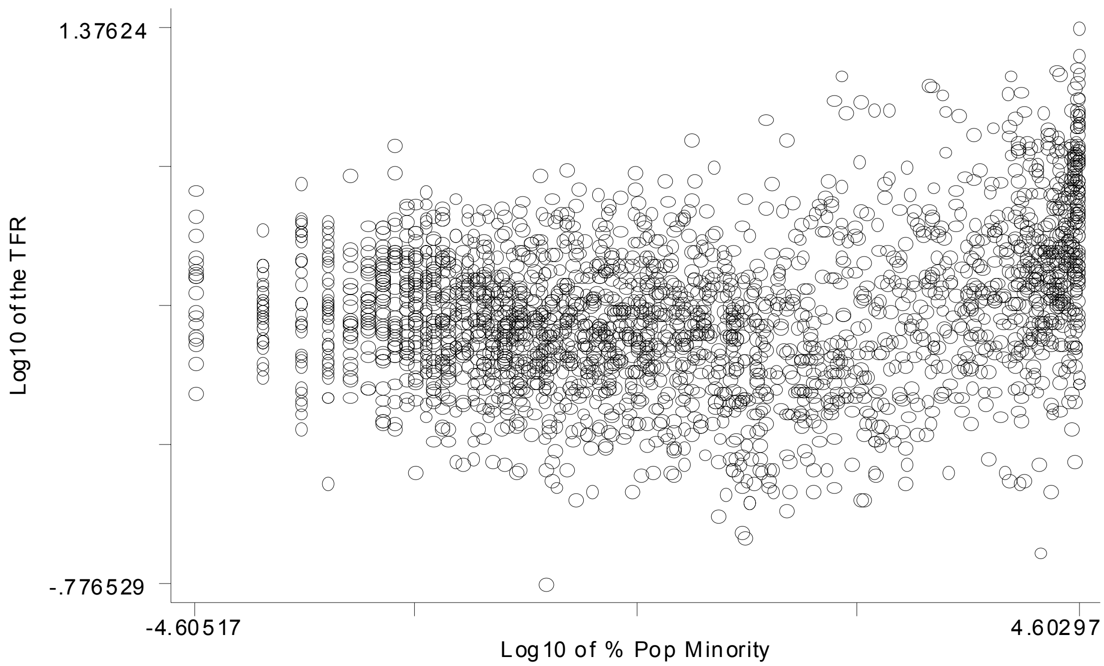

FIGURE 6: TFR & Percentage Population Minority (r = 0.232)

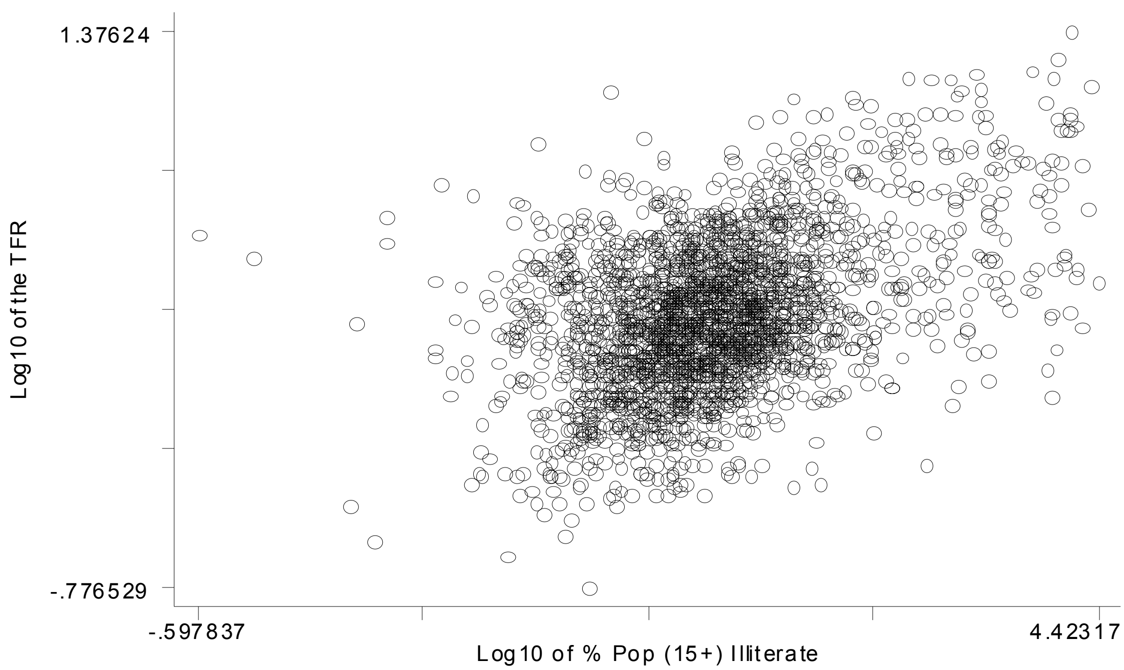

FIGURE 7: TFR & Percentage Population (15+) Illiterate (r = 0.439)

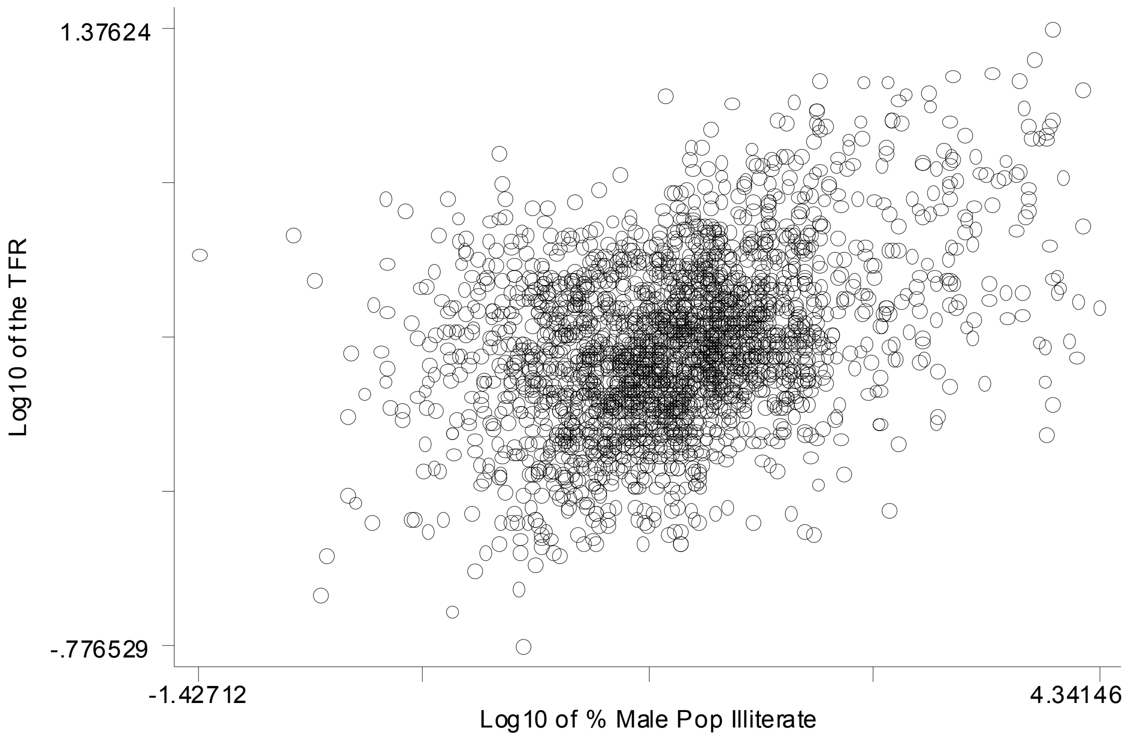

FIGURE 8: TFR & Percentage Male Population (15+) Illiterate (r = 0.394)

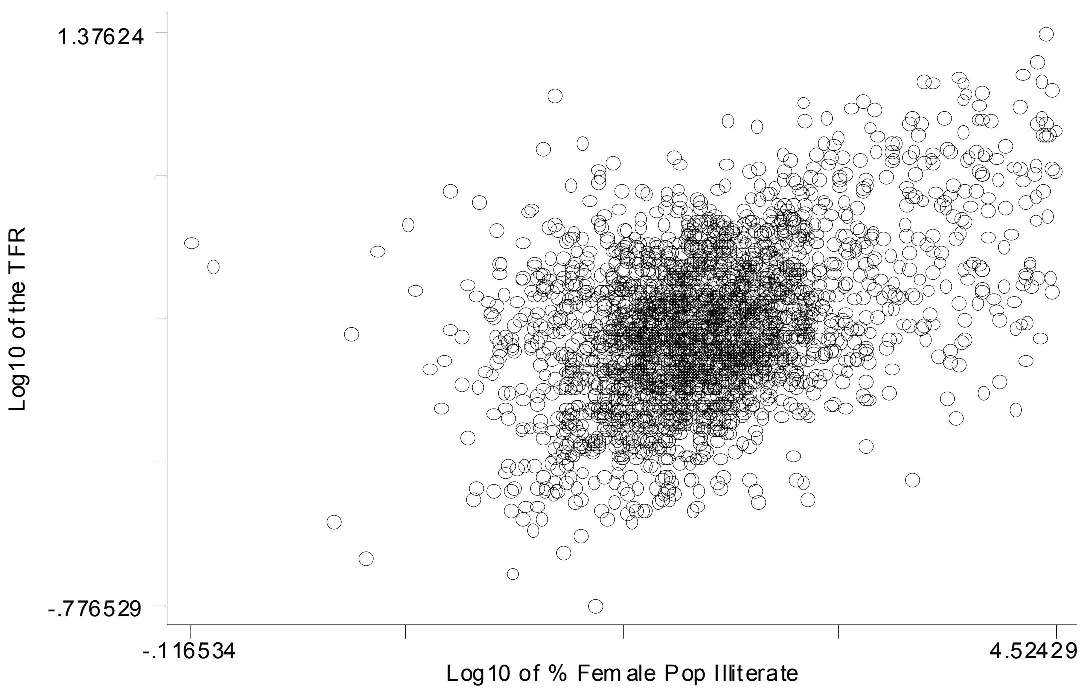

FIGURE 9: TFR & Percentage Female Population (15+) Illiterate (r = 0.441)

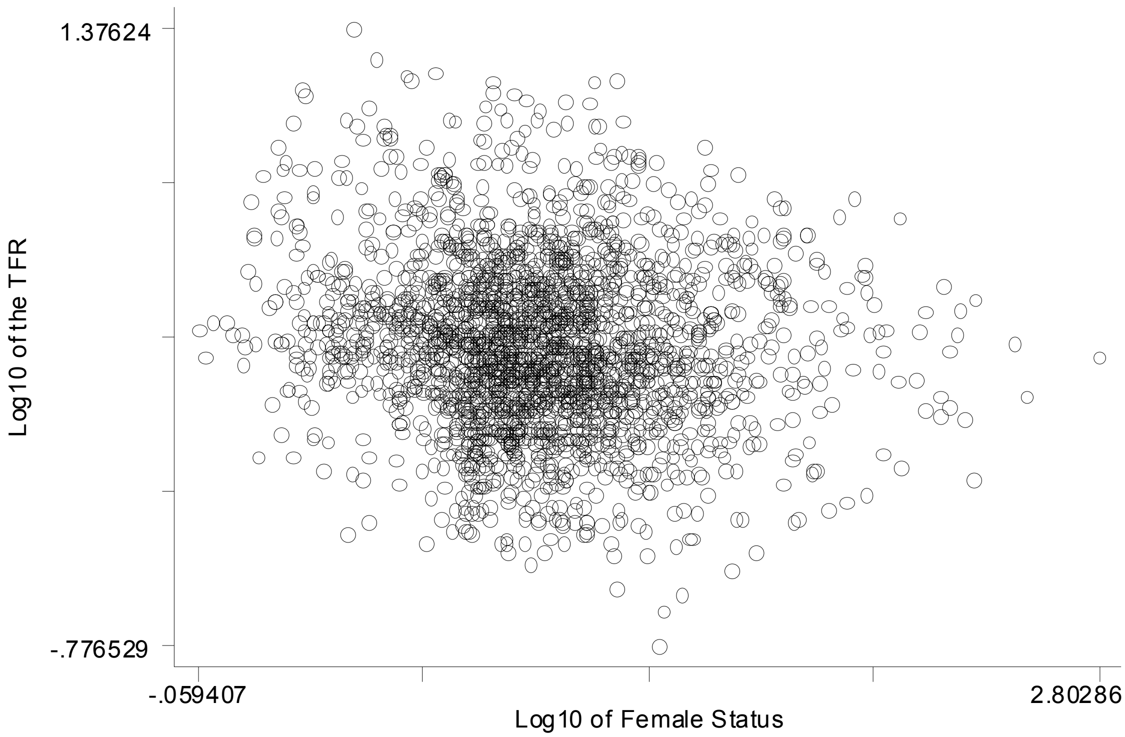

FIGURE 10: TFR and Female Status (r = -0.115)

SOURCE: 2000 Population Census of China (State Council and State Statistical Bureau 2002)

VITA

NAME: HEATHER KATHLEEN MARY TERRELL

PERMANENT ADDRESS:

1826 Co. Rd. 393

Stephenville, TX 77840

Phone: (979) 492-1249

EDUCATION BACKGROUND:

Master of Science in Sociology

Texas A&M University, College Station, Texas

May 2005

Bachelor of Science in Psychology

Texas A&M University, College Station, TX

December 2001