Title

Three Essays in Health and Development Economics

Permalink

https://escholarship.org/uc/item/9q99w0fb

Author

Xue, Yuhan

Publication Date

2014

Copyright Information

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

UNIVERSITY OF CALIFORNIA SANTA CRUZ

THREE ESSAYS IN HEALTH AND DEVELOPMENT ECONOMICS

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

by

Yuhan Xue

June 2014

The Dissertation of Yuhan Xue
is approved:

_____

Professor Carlos Dobkin, Chair

_____

Professor Justin Marion

_____

Professor Joshua Aizenman

_____

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

Abstract

Three Essays in Health and Development Economics

by

Yuhan Xue

The main focus of this dissertation is on social issues affecting people's health and welfare, within both the developed world and developing world. Particularly, policy changes are employed as natural experiments, to identify the health and social impacts of policies and welfare. This dissertation consists of three chapters.

Chapter 1 analyzes the health effects of China's centralized home heating infrastructure. The primary health mechanism is that recipients of centralized heating do not need to use traditional fuels to heat their homes, and thus suffer from less exposure to indoor air pollution. A secondary mechanism is the increased quality of modern heating leads to improvement in health. I exploit spatial and time variation in the implementation of China's centralized heating program. Spatial variation takes the form of a discrete geographical cutoff between regions that receive centralized home heating, and those that do not. I compare areas near the geographical cutoff, before and after the provision of centralized heat using a difference-in-differences framework. My outcomes of interest are perinatal mortality and prevalence of low birth weight infants, since existing research suggests this sub-population is most vulnerable to indoor air pollution. I find that areas with centralized heating have a 0.13% decrease in perinatal deaths and a decrease of 1.17% in the proportion of low birth weight infants. This translates to about 7,000 fewer deaths and about 60,000 fewer low birth weight infants each year.

Chapter 2 investigates EBT card reforms in California's food stamp program, and the impact on food insecurity. Our hypothesis is that EBT cards will reduce food insecurity by reducing the food costs associated with loss and theft of benefits, as well as decreasing fraudulent sales of benefits. We use the California Health Interview Survey, and the roll-out of EBT card reforms across California counties, to conduct an event study. Our findings suggest no evidence for a decrease in food insecurity. We do, however, find evidence of a transitory increase in food insecurity immediately following EBT card reforms. Reforms increase the likelihood of food insecurity by about 2% for 1-2 months depending on the measure of food insecurity used. The result is distinguishable from zero, and robust to changes in specification, inclusion of controls, and measurement choices. We posit the increase was due to a less than perfectly smooth transition to the EBT card system.

Chapter 3 uses large, national surveys to investigate individual attributes associated with probability of divorce. In recent years China's divorce rate has risen rapidly. With this rapid rise has come a large number of potential explanations, both those grounded in economic theory, and those widely discussed in the public discourse. We investigate which individual attributes are associated with an increased probability of divorce, and to explain which explanations are not empirically substantiated. We find that Western attitudes and a sense of relative affluence are predictive of divorce. Furthermore, most popular explanations are not empirically confirmed. Finally, we find that previous results suggesting divorce is associated with worse mental health outcomes are applicable in China and not only Western nations.

Acknowledgements

Chapter 1

Health Returns to Modern Heating: Evidence from China

1.1 Introduction

In many developing countries, use of charcoal briquettes, firewood and coal in indoor stoves are a major form of home heating during the colder months of the year. A considerable concern is that these more traditional heating methods may be injurious to human health and thus reduction in their use may be a worthwhile objective. In particular, combustion of traditional fuels creates smoke and other byproducts that are toxic to human beings. A secondary concern is that traditional heating methods tend to be highly localized and therefore less effective at keeping the home warm, thus leading to greater possibilities of respiratory problems.

In spite of the shortcomings of traditional fuels, they remain an important aspect

of home fuel usage in China. As shown in Table 1.1, more than 90% of the total residential energy consumption in China comes from traditional fuel. Given the widespread use of traditional home heating methods in China, and the tremendous numbers of people that may be adversely affected, China presents itself as one of the countries that is most heavily impacted by traditional home heating.

Table 1.1: 1992 Percentage Distribution of Houshold Energy Use

	All	Urban	Rural
Wood	31.8	0	36.4
Crop Residues	33	0	37.8
Coal	25.7	55.8	21.9
Electricity	6.3	30.4	3.1
Gas	0.7	5.7	0
Dist Heat	0.9	8.1	0
Oil Products	1.3	0	0.4
Dung	0.3	0	0.4

However, a great many Chinese are now able to heat their homes without resorting to traditional methods due to the introduction of the Chinese central heating system. The Chinese central heating system is an indoor heating system provided by the Chinese government that allows recipients to heat their homes with no need for traditional fuels. It was first introduced in the 1980's, and has been widely adopted over 17 northern provinces in China during the past 15 years, especially the past 10 years. As such, there exists considerable variation in the provision of centralized home heating over time. Under the central heating system, heat is generated by government-owned cogeneration plants and boilers, in the form of steam and hot water. After being generated, heat is then transported to houses and apartments through underground steam pipelines and hot-water pipelines. All

the houses have been built in a way that is able to accommodate the heating pipelines.

This paper studies the health impact of the Chinese central heating system. Particularly, I investigate the social health costs of burning traditional fuels indoors, using the central heating system as a natural experiment. Empirically, I use difference-in-differences (DID) as the primary means of determining the impact of central heat provision on perinatal mortality in China. Under the natural experiment setting, a simple comparison of pre-treatment and post-treatment outcomes for the individuals exposed to a treatment are likely to be affected by temporal trends in the outcome variable caused by changes in other factors which coincided with the treatment period. Thus, an untreated comparison group is used to identify temporal variation in the outcome that is not due to treatment exposure. The impact of burning traditional fuels indoors is measured by the difference before and after the introduction of the central heating system. This allows me to control for time varying differences by comparing areas that recently received central heating to similar areas that have not yet received central heating. I am also able to use spatial variation as China's central heating system has been deployed in accordance with strict geographical cut-offs that completely determine whether a city receives centralized heating or not. Thus, I am able to compare locales near the cut-off and significantly diminish concerns about omitted variable bias as these areas are generally the same near the cut-off except that one locale receives heat and another does not.

Traditional home heating methods produce byproducts of great concern including carbon monoxide, particulate matter and genotoxic substances. Carbon monoxide (CO), is a byproduct of incomplete combustion of carbon-containing compounds, and is invariably

created when plant biomass or coal is burned. Carbon monoxide is colorless, odorless, and tasteless, but highly toxic and imperceptible to human sensory organs. It combines with hemoglobin in the blood stream to produce carboxyhemoglobin, which usurps the space in hemoglobin that normally carries oxygen, thus rendering blood less effective in delivering oxygen to bodily tissues. Carbon monoxide poisoning is the most common type of fatal air poisoning in many countries. Another deleterious byproduct from the combustion of traditional fuels, are the fine particles suspended in smoke created by burning. These fine particulates are quite harmful to health, and are particularly harmful if the diameter of the particle is less than 2.5 μm. Particles smaller than 2.5 μm are transported deeply into the lungs after being inhaled. 80% to 95% of the particles in the smoke from traditional fuel combustion are smaller than 2.5 μm, with the size distribution mode 0.2 to 0.4 μm (Hueglin et al. 1997). Moreover, the smoke from traditional fuel combustion also contains genotoxic hydrocarbons such as benzo(a)pyrene, these genotoxic hydrocarbons may result in growth deficiencies and/or mental retardation in a developing fetus as they alter the DNA structure on a cellular level.[1]

An additional consideration is that traditional heating methods generally result in highly localized heating in the sense that the immediate vicinity of the stove or furnace is sufficiently heated, but the surrounding area is generally not. Thus many households utilizing traditional heating methods are, in reality, only able to heat part of their homes. More explicitly, independent of any pollution, the quality of traditional heat is generally lower and less well distributed than modern methods. As a result of poorly heated areas of the home, the very young and the very old, as well as anyone with compromised immune

---

[1]European Food Safety Authority

heath, may be left more vulnerable to infectious disease as a result of inadequate heating. In particular, research done by Carney and Klein (1965) has indicated that low birth weight infants exposed to lower temperatures face longer odds of survival as they are exposed to greater metabolic stresses. Moreover, poorly heated homes are more prone to high levels of humidity as evaporation of moisture is lower. Higher humidity in cold conditions has been associated with greater incidence of respiratory tract infections. Strachan and Sanders (1989) find evidence that damp conditions, possibly due to mold, are associated with much higher rates of childhood asthma as well.

I use early-life mortality as the main measure of health in the study. In particular, perinatal mortality is chosen as the main outcome variable in this study. I choose perinatal mortality for the following reasons. First, fetuses and infants are a population group most vulnerable to indoor air pollution. Studies have shown that exposure to pollution during the last trimester in utero is the most damaging (Jayachandran 2005). Prior research has investigated the linkages between air pollution and early-life mortality, for example Chay and Greenstone (2003) find that the passage of the Clean Air Act of 1970 led to a decline in infant deaths. A recently released paper by Chen et. al. (2014) uses an earlier Chinese heating policy to consider differences in outdoor air pollution and life expectancy. The central implication is free coal used for generating heating increased total suspended particulate matter in North China and contributed to decreased life expectancy via increased cardio-respitory disease. Currie and Neidell (2005) find that exposure to carbon monoxide and other air pollutants during the month of birth is associated with infant mortality, using 1990's data in California. Second, perinatal mortality data used in this paper is collected by

the government through hospital reports, thus it has better reliability compared to infant mortality, which is calculated indirectly from the Chinese national census. Thus, I consider the most pertinent group, with the most reliable data records.

Despite the fact that a very large number of people across the globe rely on traditional home heating and it is readily apparent that doing so may have negative health effects, researchers have generally been unable to estimate reliable measures of the health impact. This stems from the fact that measuring the social health costs of burning these fuels indoors has proven quite difficult from a variety of perspectives. One source of difficulty is that the group of individuals who burn traditional fuels on a regular basis and those who use modern heating methods generally have vey different socioeconomic backgrounds. A simple comparison between these two groups of individuals would be insufficient to engage in causal inference. One attempted solution is to control for socioeconomic background variables, such as income level and education level. However, there persists a problem with this approach in that there is likely to be considerable omitted variable bias. This stems from the fact that some key characteristics, that are correlated both with both health outcomes and burning traditional fuels indoors, are unobservable. Without controlling for these unobservable characteristics, and only controlling for observable traits, any results will suffer from omitted variable bias. Thus, ideally, experimental data are needed in order to make credible estimates of social health costs. In the absence of truly experimental data, natural experiments are widely used in empirical studies in the social sciences. Natural experiments have been used to measure health effects of air pollution in various studies of interest. Pope et al. 1992 studied the temporary closure of a steel mill in Utah during a 1986 labor dispute.

Friedman et al. 2001 studied the reduction in traffic during the 1996 Olympics in Atlanta. However, to date very little has been done with regards to indoor air pollution.

My central result is that the perinatal mortality rate for the treatment group drops about 0.13% faster than the control group over the course of 10 years. This translates to about 7000 fewer perinatal deaths as a direct consequence of centralized heating provision. That is, in the absence of centralized heating we would expect there to have been an additional 7000 perinatal deaths. As a baseline reference, perinatal mortality is about 0.6% or, 6 out of every thousand children die during the perinatal period. Another key result is that the percentage of low birth weight infants drops for the treatment group, whereas it actually increases for the control group. The difference is 1.17%. which translates to roughly 60,000 fewer low birth weight infants in the treatment group than we would otherwise expect in the absence of centralized home heating.

The rest of the paper proceeds as follows. In the next section, I provide a brief background about the central heating system in China, and review the literature on indoor air pollution and early-life mortality. The third section presents the empirical analysis of the effects of central heating system on perinatal mortality rate. The fourth section presents the empirical analysis of the effects of central heating system on the percentage of low birth weight infants. The fifth section provides further discussions in terms of model assumptions and model choice. The final section summarizes my conclusions.

1.2 Background, Data, Sample Design and Evaluation

1.2.1 Background: Traditional Home Heating

It is worth noting that while burning traditional fuels is a common social problem in developing countries, it is also a considerable health concern in the developed world. In developed countries, charcoal briquettes are used as a substitute for modern heating methods in the winter time during power outages, or when electricity has been disconnected, according to the US Consumer Product Safety Commission. Many fatal deaths have occurred, even within the US. Over 170 people in the United States die every year from carbon monoxide produced by non-automotive consumer products. Hampson et. al (1994) studied cases of patients with unintentional CO poisoning from 10 counties within the state of Washington between October 1982 and October 1993. Of the 509 patients treated for acute unintentional CO poisoning, 79 cases occurred in 32 incidents as a result of indoor burning of charcoal briquets, for the purpose of either home heating or cooking. Thus the issue should not only be thought of as relevant to the developing world. However, as is expected the majority of ill effects are to be found in the developing world.

In general, home heating stoves and fuels in China are quite heterogenous with some discernible patterns. Primary fuels are coal, charcoal briquettes, crop residues and firewood. Firewood and crop residues are primarily restricted to rural settings as the fuel itself occupies a considerable amount of space and is quite heavy. Moreover, typically only those households engaged in grain cultivation have access to large quantities of crop residues. In the case of urban Chinese, they typically live in small apartments (by Western standards) and frequently not at ground level, but rather up several flights of stairs. Thus

firewood is very unpopular in urban settings. Coal is quite popular in that China has large
and extensive reserves of coal, particularly in northern China. Charcoal is also very popular
and particularly in areas where coal is less prevalent. Presumably, coal would be weakly
preferred since it requires less storage space. Coal and charcoal are roughly equivalent in
price (Table 1.2).

Table 1.2: Estimated Costs of Various Heating Methods

Heating Method	Unit Price	Heating Usage	Annual Cost
Charcoal Briquettes and Coal	$0.1 per kilogram	10kg per day	$150
Firewood	$50 per cart	5-10kg per day	$200
Central Heating	$4 per sq. meter	60 sq. meters	$240

Stoves are also highly heterogenous across China, not only in terms of the fuel
they are intended to be used with, but also in terms of size and ventilation properties. In
many parts of northern China a type of stove known as a *kang* predominates. A *kang* is
essentially a large raised platform with a single large piece of masonry that is used to retain
heat for long periods of time. Traditionally the *kang* is used as a sleeping surface and also
as the preferred location for most day time activities. The degree to which the *kang* is
enclosed or vented varies tremendously. The *kang* is typically quite large and may occupy
a third or half of a living room. Typically the *kang* is constructed so that a portion of
it protrudes into an adjacent kitchen and can be used for cooking purposes. Other stove
types are more in line with typical conceptions of Western stoves. The most primitive of
stoves (but quite common) are essentially a large cauldron on legs. Typically these stoves
are only used for coal and charcoal as they are not able to ventilate large amounts of smoke,
and thus their use with firewood or crop residues would be impractical (the room would be

absolutely flooded with smoke).

Ventilation is also quite heterogenous in that some stoves are connected to chimneys and flues that vent smoke out of the building, while others do not. Furthermore, the general level of ventilation efficacy for ventilated stoves is highly varied. Cauldron type stoves invariably raise indoor pollution levels substantially, while the best ventilated and modern stoves may have much more minor effects on indoor pollution. A relevant and common practice is to heat coal or charcoal in a stove and once it is lit to transfer the lit fuel to other parts of the house where there is no ventilation system in place to draw away the less visible, but still highly present smoke and combustion byproducts.

A final, but important consideration, is to recognize that in addition to problems associated with indoor air pollution, traditional methods suffer from another considerable shortcoming. Traditional heating methods tend to be fairly ineffective from the perspective of diffusing heat throughout a dwelling in an even and well distributed manner. It is extremely common to find that the immediate vicinity around a stove is quite warm, while the other side of the room is quite cold. Furthermore, other rooms in the home are likely to be even colder. Cold and damp conditions are generally recognized as compromising immune health and increasing the risk of infectious disease, particularly respiratory ailments. Thus, access to modern heating methods may improve health outcomes via reduced indoor air pollution, but also via improved ambient temperature and humidity. Humidity is affected in that poorly heated homes tend to be far damper due to diminished evaporation of moisture in the home. This elevated humidity may be injurious in itself, but unequivocally leads to higher growth of molds, which have been clearly demonstrated to adversely affect respiratory

health.

1.2.2 Background: Central Heating System

I next provide a brief synopsis of the central heating system in China. The central heating system has only been introduced to the northern half of China, as there exists large temperature differences between northern China and southern China. Within China, there is a clear and widely accepted geographical line that serves as the traditional North-South divide - the *Tsinling-Huaihe* line. *Tsinling* refers to a major mountain range in central China. It stretches from Gansu and Shaanxi Provinces in the west to Hebei and Henan Provinces further east. Just as the *Tsinling* range diminishes to low foothills the *Huaihe* river emerges from the eastern edge of the mountains. The *Huaihe* is one of the chief rivers in China, it generally borders Hubei, Henan, Anhui, Shandong and Jiangsu Provinces. Thus, these two geographical features have traditionally served as a clear and intuitive divide between North China and South China. It should be noted that areas to the west of *Tsinling* mountains are far too low in population density to merit central heating. The central heating system has as at present, been built only in provinces to the north of the *Tsinling-Huaihe* line. While no provinces to the south have yet received government provided central heating. Figure 1.1 shows the rough location of *Tsinling-Huaihe* line on a map of China. It should be noted that many provinces that are just to the south of the *Tsinling-Huaihe* line still have cold, harsh, and snowy winters that necessitate indoor heating, but the central heating system has not yet been built in these provinces. It is generally unknown within the public domain if the Chinese government plans to extend heating infrastructure to the South anytime soon.

Figure 1.1: Map of China: *Tsinling-Huaihe* Cutoff

Over the past 15 year, with the introduction of the central heating system, most households in northern China have switched from charcoal or coal burning to central heating as their major indoor heating method (Table 1.1). The central heating system provides consistent and reliable heat without adverse pollution of indoor air. Moreover, it provides heat that is well distributed throughout the home. In particular, the central heating system does not allow for the possibility of heating related carbon monoxide, or particulate matter to enter the home as a consequence of traditional home heating. A peculiarity of the central heating system is that the government-owned cogeneration plants and boilers only operate during a fixed period of a year to provide heat to nearby houses and apartments. For most cities, the heating period is from November to March. However, there is some year-to-year variation (for example in 2012 the "heating season" started a few weeks early). Thus, there

persists the possibility of traditional home heating for brief periods right before, and after the "heating season". Throughout the "heating season", houses and apartments under the central heating system are heated 24 hours per day, with an indoor temperature of 16-18C (60.8-64.4F). While the households in northern China have switched to central heating, households in southern provinces with cold winters continue to rely on coal and charcoal burning as the major indoor heating method.

I use central heating system data to identify the health effect from home heating via traditional indoor heating methods. The choice of using the central heating system as the treatment is motivated by the following considerations. First, under the central heating system, the choice of whether a city is to be centrally heated is exogenous to all population characteristics and is determined solely by a largely arbitrary line. That is, if a city is on the northern or southern side of the *Tsinling-Huaihe* cutoff. More explicitly, whether a city receives the treatment is not affected by economic background or by some measure of health status within the city. Thus, there is no endogenous placement of centralized heating. Second, the central heating system was introduced in the 1990's and has expanded at a rapid pace during the past 15 years resulting in considerable variation in heat exposure over time. Figure 1.2 shows the heated area per person for the past 15 years in northern China. Note the considerable increase over time attributable to the expansion of central heating. Heated area per person for southern China has remained relatively constant over time, and thus allows for Southerners just south of the cutoff to serve as a control group. In northern China, heated area per person increases from 1.66 square meters in 1996 to 10.02 square meters per Person in 2011. Between the pre-treatment time period and the post-

treatment period, there is considerable variation within the sample. A third consideration is that both the treated and untreated populations are very large and as a result estimates are fairly precise. The central heating system covers 17 provinces and 0.56 billion people [2], which provides a rich sample size for research and analysis.

The central heating system data are from *China Statistics Yearbook* and *China Urban Construction Statistics Yearbook*. Data are available at the city level for all cities where the central heating system has been introduced. For nearly all cities (and an overwhelming majority of cities along the cutoff), data can be traced back to the early 1990's.

Figure 1.2: Heated Area Per Person
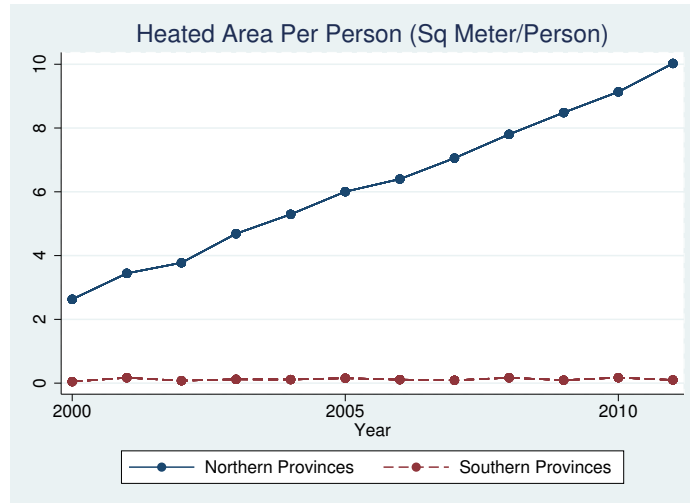
1.2.3 Data

A considerable challenge in using the central heating system is the issue of data availability. Data availability in China is generally much more limited than in the US or Europe, this is especially true of micro level data, which are generally unavailable. Central heating system data are primarily from the *China Urban Construction Statistics Yearbook*,

---

[2] 2011 population data, *China Statistics Yearbook* 2012

and perinatal mortality data are from the *China Health Statistics Yearbook*. Data from both yearbooks are collected by the Chinese government on an annual basis. One potential disadvantage of using yearbook data is, the sample size is relatively smaller than micro level data, which makes the data less powerful in terms of statistical inference. Here I adopt Abadie, Diamond, and Hainmueller's (2010) argument that there is no estimation uncertainty at the aggregate level, since the goal is to determine the effect of a policy on an entire population rather than a particular sub-population.

1.2.4   Background: Perinatal Mortality and Low Birth Weight Infants

Perinatal mortality is chosen to be the main outcome variable in this paper. The perinatal period is defined as the period starting from 28th week of gestation and ending 4 weeks after birth. I choose perinatal mortality since air pollution is most damaging to fetuses and extremely young infants. The period of peak vulnerability is believed to be during the last trimester of pregnancy. Various research projects have studied the impacts and the mechanism by which exposure to air pollution causes damage to fetuses and infants. Chief among them are Dejmek et al. 1999, Wang et al. 1997, Berkowitz et al. 2003.

Central among established findings are that during the 4 weeks following birth, postnatal exposure to carbon monoxide at even fairly low levels can lead to acute carbon monoxide poisoning. Moreover, exposure to air pollution can lead to acute respiratory infections, one of the leading causes of infant mortality. Prenatal exposure to air pollution is believed to affect fetal development in the following ways. First, the mother's exposure to air pollution interferes with the mother's health, thus fetal nutrition and blood flow is disrupted. Furthermore, after the mother is exposed to air pollution, toxic substances

15

enter into mother's bloodstream and cross the placenta, thus affecting the developing fetus. Depending on the toxicants, fetal growth deficiency, retardation, low birth weight, low head circumference and/or shorter gestational periods may result as a consequence of air pollution.

For perinatal mortality, I use data from *China Health Statistics Yearbook*. Perinatal mortality are available at the province level from 2002 to 2011. Figure 1.3 shows the trend of average perinatal mortality in China.

Figure 1.3: Perinatal Mortality in China, 2002-2011
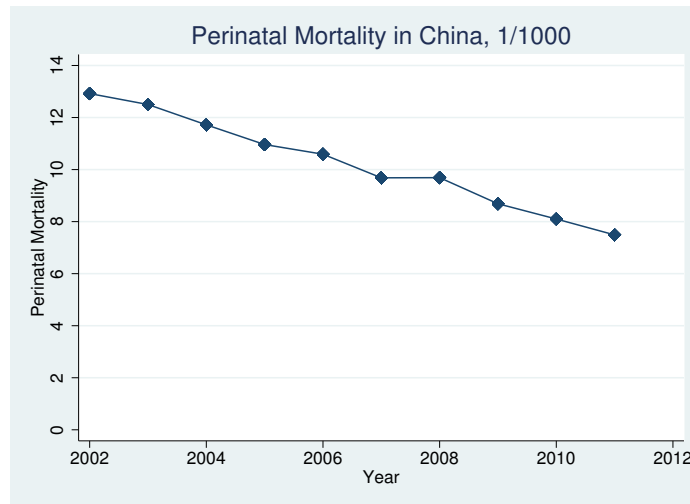
In addition to perinatal mortality, I also consider the proportion of infants born with low birth weight. This metric is a widely used and accepted proxy for the overall health and development of a newborn infant. While this measure does little to inform us of any root causes, it provides a reliable and readily understood indicator for babies that have endured a difficult gestational period. Furthermore, as birth weight is recorded for all babies born in hospitals, it represents a data source that is likely to be highly reliable. Also inclusion of infant birth weight allows me to confirm that my results relating to perinatal

mortality are roughly mirrored by a measure that we would expect to be highly correlated.

Existing research suggests the existence of considerable evidence of a fairly robust correlation between exposure to biomass smoke and infants born with low birth weights. Recent studies have found this linkage to be related to fetal growth retardation and/or shorter gestational periods. Both fetal growth retardation and shorter gestational periods are linked to maternal exposure to outdoor air pollution at levels of pollution substantially lower than what is typically found in biomass-burning homes (Dejmek et al. 1999, Ritz et al. 1999, Chen et al. 2002). Thus, it is highly likely that indoor pollution in homes using biomass for heating is more than sufficient to be associated with a reduction in birth weight for many infants. Besides exposure to deleterious particulate material, there is the risk of exposure to carbon monoxide, which has been associated with suboptimal fetal development and adverse pregnancy outcomes, including reduced birth weight (Astrup et al.1972). Levels of CO in homes using biomass fuels are sometimes high enough to result in carboxyhemoglobin levels comparable to those in smokers (Dary et al. 1981, Behera et al. 1988). For research based on data from developing countries, Boy et al. (2002) used a sample of children in rural Guatemala and found an association between household use of wood fuels and reduced birth weight. Mishra et al.(2004) found similar results in Zimbabwe.

1.3 Estimation using Perinatal Mortality Rates

1.3.1 Baseline Difference-in-differences Model

The goal of the empirical analysis is to examine whether burning traditional fuels indoors has an effect on perinatal death. Under the standard, or baseline, difference-in-

17

differences model, outcomes are observed for two groups for two time periods. One of the groups is exposed to a treatment in the second period, but not in the first period. The second group is not exposed to the treatment during either period. In the case where the same units within a group are observed in each time period, the average gain in the second (control) group is subtracted from the average gain in the first(treatment) group. This removes biases in second period comparisons between the treatment and control group that could be the result from permanent differences between those groups, as well as biases from comparisons over time in the treatment group that could be the result of differential trends. In order for the approach to uncover credible estimates of the impact of traditional fuel home heating on perinatal health impacts, it must be the case that the treatment (provision of central heat) is independent of all observable and unobservable characteristics of the treatment and control groups.

In this study, treatment status is solely determined by the geographic divide between treatment and control group, i.e. *Tsinling-Huaihe* line. The treatment group and control groups are determined on the basis of their location relative to this geographic cut-off. Particularly, the treatment group consists of provinces that are on the northern side of the *Tsinling-Huaihe* line, and are close to the *Tsinling-Huaihe* line. Provinces farther away from the *Tsinling-Huaihe* line are more likely to have considerable differences, both observable and unobservable, with the control group in this study. Thus those provinces well north of the *Tsinling-Huaihe* line are omitted from the empirical analysis. There are 10 provinces in the treatment group, namely Shandong, Ningxia, Tianjin, Beijing, Gansu, Henan, Shaanxi, Shanxi, Qinghai, and Hebei Provinces. The control group consists of

Table 1.3: Comparison Between Treatment and Control Group

Group	Treatment	Control
Population(10,000)	4074.65	5624.92
UrbanPopulation(10,000)	1994.00	2999.35
Percentage of Urban Population	53.39	56.51
GDP2011	15691.07	23283.16
Illiteracy Rate	5.7	5.52875
Gender Ratio	105.143	104.06625
2010 Life Expectancy	75.02	76.21
Birth Rate(1/1000)	11.333	10.20875
General Smoking Rate(%)	36.5	37.93
Smoking Intensity(No. of Cigarettes)	14.34	15.39

provinces that are on the southern side of the *Tsinling-Huaihe* line. They either explicitly border the line, or are very close (within 150 miles). Each province in the control group has a mean January temperature below 35F, and has low temperatures well below freezing.[3] There are 8 provinces in the control group, namely Anhui, Chongqing, Hunan, Hubei, Jiangsu, Shanghai, Sichuan, and Zhejiang Provinces. Provinces farther south than those in the control group have higher temperatures in the winter months, therefore people rely less on indoor heating in these provinces. Moreover, as one moves further from the cutoff, the assumption of the treatment status being orthogonal to observable and unobservable characteristics becomes tenuous.

Table 1.3 displays a brief list of characteristics between the treatment and control group. The two groups have similar backgrounds in terms of economic development and people's behavior and habits.

---

[3] *China Atmosphere Yearbook.* Data are average temperatures over the past 10 years.

In the baseline difference-in-differences model, Let

$$PMR_1 = \text{Perinatal mortality for treated provinces}$$

$$PMR_0 = \text{Perinatal mortality for control provinces}$$

be the counterfactual outcomes. It is reasonable to assume the conditional mean independence in this case:

$$E[PMR_0|D = 1, X] = E[PMR_0|D = 0, X] = E[PMR_0|X]$$

i.e. perinatal mortality without central heating system, y_0, does not determine participation. That is to say, the Chinese government has determined provision of centralized heating in a way that does not depend on perinatal mortality patterns.

Here X takes the particular functional form

$$E[PMR_0|X] = E[PMR_0|g, t] = \gamma_g + \lambda_t$$

where g denotes the group (treatment or control) and t denotes the period. In a baseline difference-in-differences model, 2 periods are considered: before treatment and after treatment. In this case the first year (2002) is used as the before treatment period, and the last year (2011) is used as the after treatment period. The difference-in-differences estimator β in this case is obtained by first taking conditional expectations for different groups and

periods, then subtracting each other:

$$\beta_{BaselineDID} = (E[PMR_i|g = \text{Treated}, t = 2011] - E[PMR_i|g = \text{Treated}, t = 2002])$$

$$- (E[PMR_i|g = \text{Control}, t = 2011] - E[PMR_i|g = \text{Control}, t = 2002])$$

Table 1.4 displays the estimates. Two sets of difference-in-differences are estimated here. For the first three columns, both perinatal mortality rate and heated area per person are calculated by averaging among provinces, with an equal weight for each province. For the second three columns, these averages are weighted by population of each province. Since provinces vary largely in terms of population, weighting by population provides a more precise estimate of any treatment effects. For the first estimation, row (1) and (2) shows the treatment group and control group, row (3) summarizes the differences. Columns (1) and (2) show the pretreatment outcomes and post-treatment outcomes, column (3) summarizes the differences. The model estimate of the baseline difference-in-differences model is 1.506, that is the effect of replacing traditional indoor heating with central heating during the past 10 years. Like other two-by-two difference-in-differences models, this model is just-identified, therefore statistical inference does not apply to this model. For the second estimation, row (4) and (5) shows the treatment group and control group, row (6) summarizes the differences. Columns (1) and (2) show the pre-treatment outcomes and post-treatment outcomes, column (3) summarizes the differences. The effect of replacing traditional indoor heating with central heating is estimated to be 1.255 under the weighted two-by-two model. This model is also just-identified, thus statistical tests are not applicable.

Table 1.4: Baseline Difference-in-differences Model

	Original			Weighted		
	(1) 2002	(2) 2011	(3) Difference	(4) 2002	(5) 2011	(6) Difference
Treat	13.898	7.458	6.440	12.374	5.941	6.433
	(5.656)	(2.694)		–	–	
Control	9.873	4.939	4.934	10.329	5.151	5.178
	(3.262)	(1.178)		–	–	
Difference	4.025	2.519	1.506	2.045	0.790	1.255

1.3.2 Regression DID with Multiple Time Periods

As the data contains 10 years of observations, to incorporate all the information substantially increases the efficiency of the estimation. The regression formulation of the difference-in-differences model is used to incorporate additional periods in the analysis. The difference-in-differences evaluation is equivalent to estimating the following regression equation:

$$PMR_{gt} = \gamma_g + \lambda_t + \beta M_{gt} + \epsilon_{gt} \tag{1.1}$$

Where the variable M_{gt} is a measure of the amount of treatment received in group g at time t, in this case, heated area per person under the central heating system. Also, under the regression formulation, it is possible to introduce controls for other factors X_{gt} that might affect the trends at the group level, such as GDP growth rates and outdoor air quality.

$$PMR_{gt} = \gamma_g + \lambda_t + \beta M_{gt} + \delta X_{gt} + \epsilon_{gt} \tag{1.2}$$

Note that estimating the difference-in-differences model using Equation 1.2 is equivalent to estimating a two-way fixed effect model. Table 1.5 displays the results. Three sets of estimation results are displayed. For the first three columns, columns (1)-(3), both dependent variable and independent variables are calculated by averaging among provinces, with an equal weight for each province. Fully robust estimation and inference are applied, as the effect is estimated at the group level. The standard error is calculated using standard errors clustered at the group level. For the second and third set of estimation results, Columns(4)-(6), and columns (7)-(9), these averages are weighted by the population of each province, taking into account the population variation across provinces. Columns (7)-(9) use robust estimation, whereas standard estimation results are shown in columns (4)-(6), for completeness. For the first column of each set of estimation, namely columns (1), (4), and (7), the effect is estimated without controlling for group specific covariates. For the second and third columns of each set of estimation, group specific factors are considered as they might potentially affect the parallel trend across groups. The second column of each set of estimation, namely columns (2), (5), and (8), incorporates the GDP growth rates between the treatment and control group. The third column of each set of estimates, namely columns (3), (6), and (9), incorporates the outdoor air quality between the treatment and control group. Outdoor air quality is measured by averaging the days in one year that the overall outdoor air quality qualifies as second degree polluted or better by Chinese government standards. I find that increasing heated area per person reduces the perinatal mortality rate. These results are statistically significant, consistent and robust across all estimation frameworks. On average, every square meter increase of centrally heated area per person

leads to a decrease in perinatal mortality rate by 0.22-0.25 ‰.

Table 1.5: Regression Difference-in-Differences Model

	FE Robust			FE Weighted			FE Weighted Robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Heated Area per Person	-0.340*	-0.355*	-0.349*	-0.223**	-0.241**	-0.250**	-0.223*	-0.241*	-0.250*
	(-0.006)	(0.022)	(0.018)	(0.059)	(0.057)	(0.085)	(0.009)	(0.018)	(0.015)
Year	-0.542**	-0.530*	-0.471	-0.580**	-0.566**	-0.570**	-0.580**	-0.566*	-0.570*
	(0.004)	(0.017)	(0.041)	(0.023)	(0.023)	(0.059)	(0.006)	(0.013)	(0.042)
GDP Growth Rate		-0.060			-0.075			-0.075	
		(0.056)			(0.043)			(0.025)	
Outdoor Air Quality			-0.0096			-0.0006			-0.0006
			(0.0066)			(0.0091)			(0.0063)

Figure 1.4 displays the trends for both treatment and control group. It is clear in the figure that the treatment group experienced a greater decline in terms of perinatal mortality rate.

Figure 1.4: Perinatal Mortality Rate
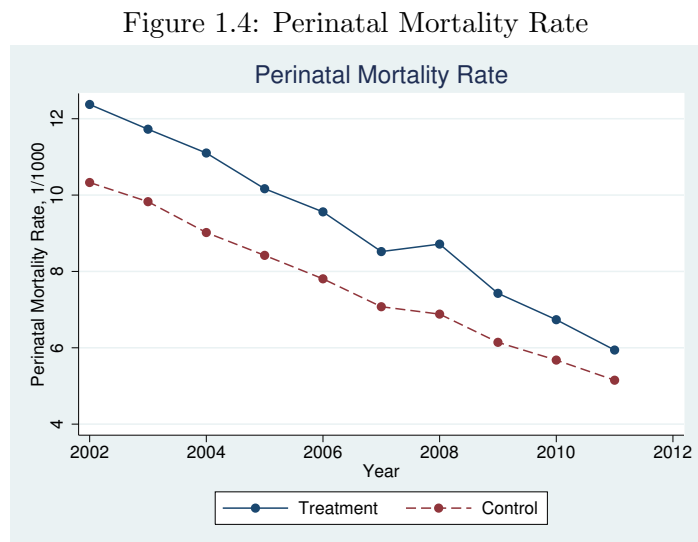
1.4 Estimation using Percentage of Low Birth Weight Infants

1.4.1 Baseline Difference-in-differences Model

This paper starts from the two-by-two difference-in-differences model. Compared with the regression difference-in-differences model, the two-by-two DID provides a baseline measurement of the treatment effect without assuming linearity. The two-by-two DID estimator is:

$$\beta_{BaselineDID} = (E[LBW_i|g = \text{Treated}, t = 2011] - E[LBW_i|g = \text{Treated}, t = 2002])$$

$$- (E[LBW_i|g = \text{Control}, t = 2011] - E[LBW_i|g = \text{Control}, t = 2002])$$

where LBW_1 and LBW_0 are the percentage of low birth weight infants in the treatment group and control group, respectively. Table 1.6 provides the estimation results of the two-by-two model. Two sets of estimates are presented. For columns (1)-(3), both perinatal mortality rate and heated area per person are calculated by averaging across provinces, with an equal weight for each province. For Columns (4)-(6), these averages are weighted by the population of each province to capture the inter-provincial variation. Columns (1) and (4) show the pre-treatment outcomes, columns (2) and (5) show the post-treatment outcomes, column (3) and (6) summarizes the differences. For both sets of estimates, the first row displays the percentage of infants with birth weight less than 2500 grams in the treatment group. The second row displays the percentage of infants with birth weight less than 2500 grams in the control group. In either set of estimates, the treatment group saw a sizable decrease in percentage of low birth weight infants between 2002-2011, whereas the control group, in contrast, had an increase. The difference-in-differences estimate is obtained by differencing the two changes. In the weighted case, the estimate is 1.171, i.e. for every square meter increase in centrally heated area per person, the percentage of low birth weight infants drops by 1.171 percent. This model is also just-identified thus statistical inference cannot be carried out.

1.4.2 Regression Difference-in-Differences Model

I use a regression difference-in-differences model to further evaluate the treatment effect. As multiple years of data are available, use of a regression difference-in-differences model increases the efficiency of the estimation and the power of inference relative to the two-by-two model. The regression difference-in-differences model is equivalent to estimating

Table 1.6: Baseline Difference-in-differences Model, Percentage of Low Birth Weight Infants

	Original			Weighted		
	(1)	(2)	(3)	(4)	(5)	(6)
	2002	2011	Difference	2002	2011	Difference
Treat	2.838	2.410	0.428	3.013	2.087	0.926
	(1.413)	(0.794)		–	–	
Control	1.855	2.083	-0.228	1.739	1.984	-0.245
	(0.525)	(0.878)		–	–	
Difference	0.983	0.327	0.656	1.274	0.103	1.171

the following fixed effects model:

$$LBW_{gt} = \gamma_g + \lambda_t + \beta M_{gt} + \xi X_{gt} + \epsilon_{gt} \tag{1.3}$$

Where the variable M_{gt} is heated area per person under the central heating system in group g at time t. Control variables X_{gt} are also introduced to control for other factors that might affect the trends at the group level, such as GDP growth rates and outdoor air quality. The regression equation is estimated using fixed effects, and the results are displayed in Table 1.7. Three sets of regression results are presented. For the first three columns, columns (1)-(3), both dependent variable and independent variables are calculated by averaging among provinces, with an equal weight for each province. Fully robust estimation and inference are applied, as the effect is estimated at group level. The standard error is calculated using standard errors clustered at the group level. For the second and third set of estimation results, columns (4)-(6), and columns (7)-(9), these averages are weighted by the population of each province, thus accounting for population variation across provinces. Columns (7)-(9) use robust estimation, whereas standard estimation results are shown in columns (4)-(6), for

completeness. For the first column of each set of estimation, namely columns (1), (4), and (7), the effect is estimated without controlling for group specific covariates. For the second and third columns of each set of estimation, group specific factors are considered as they might potentially affect the parallel trend across groups. The second column of each set of estimation, namely columns (2), (5), and (8), incorporates the GDP growth rates between the treatment and control group. The third column of each set of regressions, namely columns (3), (6), and (9), incorporates the outdoor air quality between the treatment and control group. I find that increasing heated area per person reduces the percentage of low birth weight infants, these results are statistically significant, consistent and robust across all specifications. On average, every square meter increase of heated area per person leads to a decrease of low birth weight infants by 0.23-0.24 %.

Table 1.7: Regression Difference-in-Differences Model, Percentage of Low Birth Weight Infants

	FE Robust			FE Weighted			FE Weighted Robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Heated Area Per Person	-.1406*	-.1524*	-.1281*	-.2329**	-.2450**	-.2398**	-.2329**	-.2450**	-.2398*
	(.0028)	(.0028)	(.0061)	(.0231)	(.0171)	(.0232)	(.0019)	(.0011)	(.0045)
Year	.0405*	.0502*	.0724*	.0406**	.0505**	.0630**	.0406*	.0505**	.0630**
	(.0022)	(.0027)	(.0027)	(.0090)	(.0070)	(.0153)	(.0013)	(.0006)	(.00007)
GDP Growth Rate		-.0496	-.0223*		-.0506**	-.0355*		-.0506	-.03555
		(.0080)	(.0003)		(.0130)	(.0162)		(.0046)	(.0055)
Outdoor Air Quality			-.0035			-.0015			-.0015
			(.0007)			(.0023)			(.0002)

Figure 1.5 displays the trends for both treatment group and control group. It is clearly shown in the figure that the treatment group experienced a greater decline in low birth weight infants.

Figure 1.5: Perinatal Mortality, Treated vs. Control
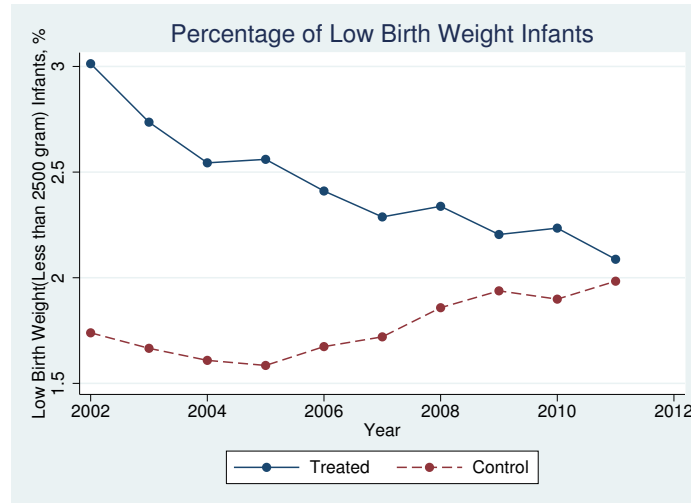
1.5 Discussion

1.5.1 Assumption Revisited: Parallel Trend

Difference-in-differences is one of the preferred methods in making causal inference due to its applicability in quasi-experimental or natural experiment settings. By measuring the difference between the pre-post, within-subjects differences of the treatment and control groups, the DID estimator provides more precise estimates compared to a simple within-subjects estimate of the treatment effect, or a between-subjects estimate of the treatment effect. However, the DID estimator holds only if all the assumptions are satisfied.

DID requires a parallel trend assumption, in addition to the standard OLS assumptions. The parallel trend assumption says that, had there been no treatment, the

trend of the outcome variable would have been the same in both treatment and control group. Under the parallel trend assumption, the only source of change between the pre-post, within-subjects differences of the treatment and control groups, is the treatment effect. A common shortcoming of many identification strategies occurs when something other than the treatment changes in one group, but not the other, at the same time as the treatment. This would imply a violation of the parallel trend assumption. In this paper, the parallel trend assumption is that perinatal mortality trends, and low birth weight infant trends, would have been the same in both groups in the absence of the treatment. Hence, the levels in the two groups can differ, but the trend in the treatment group has the same slope as in the control group. The parallel trend assumption is difficult to prove or verify, but the following ways provide some potential solutions when considering the issue.

In general, it is much preferred to have relatively smooth trends. In the case of volatile trends, any treatment effect may be largely confounded by the inherently noisy nature of the data. Card and Krueger (2000) examined time series of administrative payroll data for restaurants in New Jersey and Pennsylvania. The data show a large amount of ups and downs in employment in the two states over several years. Thus the correlation between minimum wage change and employment rate movement in their former work (Card and Krueger 1994) might not be causal, instead, it is potentially likely to be part of the random noise. However, in this paper, random noise appears quite minimal as the trend is quite smooth across both treatment and control. More specifically, there is a fairly smooth decline in perinatal mortality for both groups that appear to have nearly perfect comovement as shown in Figure 1.4 and Figure 1.5. This is true over the 10 years of data,

besides the fact that perinatal mortality declines at an increasing rate in the treatment groups, the trend between the treatment group and control group is also smooth and highly consistent. Moreover, there is a small kink in 2008 in both treatment and control group, which suggests quite strongly that the control group mimics the treatment group even to the extent of showing what appears to be an aberrant kink in 2008. On a simpler level, this also suggests that the control group is in fact a good control group in that it provides a highly relevant counterfactual.

Another consideration is that there may be factors that are changing in one group, but not the other, at the same time as the treatment. However, if these shortcomings are identified and controlled for, we are restored to a valid assumption of parallel trends. Thus, violations of the parallel trend assumption may not prove to be injurious if the only violations are observable and can be corrected for. By controlling for GDP growth rate changes, and outdoor air quality changes, in the regression difference-in-differences model I account for two potentially problematic differences between the treatment and control groups. That is, in the absence of these controls, it is possible that my treatment effect would be driven not by centralized home heating, but rather by changes in outdoor pollution or changes in the overall level of economic development. Furthermore, it should be noted that these particular concerns (differential outdoor pollution and differential economic growth rates) are issues that are likely to influence perinatal mortality and infant birth weight in a direct and meaningful way. Figure 1.6 displays the trends of GDP growth rates and outdoor air quality for both treatment and control groups. Despite the obvious need for care and attention, these factors do not differ systematically between the two groups, and

importantly, neither are statistically significant in regression models. Reassuringly, the central results appear to be quite robust to the inclusion of these controls.

Figure 1.6: GDP Growth Rates and Outdoor Air Quality
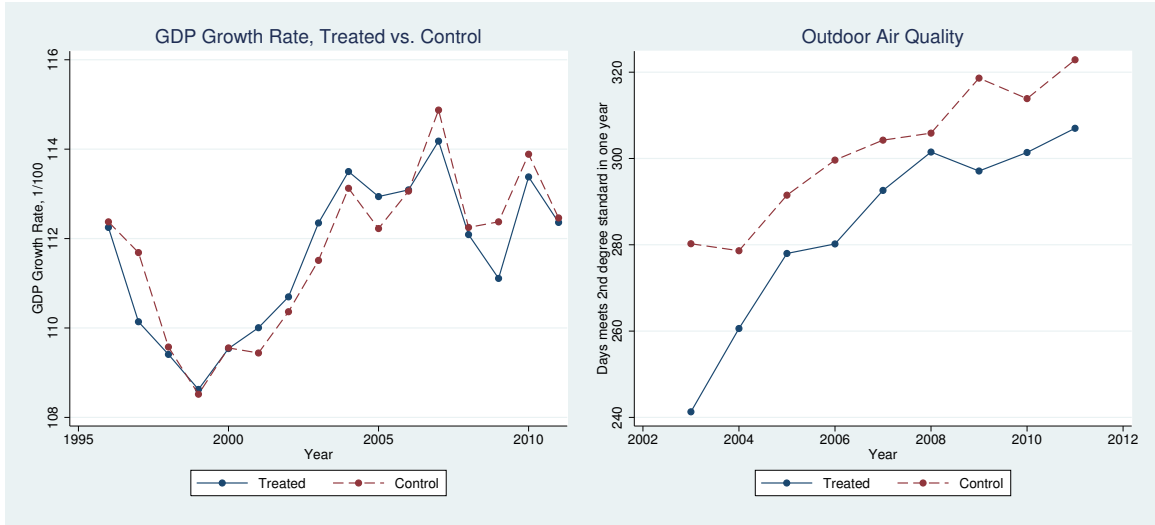
A third consideration is the pre-treatment trends.If the control group has similar pre-treatment trends to the treatment group, then the finding that a treatment effect does indeed exist is far more convincing. Literature using this method include Autor (2003) and Hastings (2004). In this context, it is not possible to obtain pre-treatment data in terms of perinatal mortality rates or low birth weight infants, due to the very limited nature of data availability in China.[4] Therefore, pre-treatment trends cannot be evaluated directly in the strictest sense. However, I am able to find a desirable proxy in the form of pre-treatment infant mortality rates. The infant mortality rates should provide an indirect measure of the pre-treatment trend in perinatal mortality in-so-much as perinatal mortality tracks infant mortality. The infant mortality statistics are from the 1990 census and the 2000 census. Results are shown in Table 1.8, note that the infant mortality rate displays similar trends

---

[4]Chinese government did not start publishing Health Statistics Yearbooks since 2003.

Table 1.8: Infant Mortality Rates for Pre-Treatment Periods

	Treat	Control	Difference
1990	31.405	26.305	5.080
2000	21.073	17.105	3.968
Difference	10.332	9.220	1.112

in both treatment and control groups. It should be noted that the two trends are a near mirror image of one another in terms of slope over time. Thus, there appears to be no evidence for a time-varying difference that would suggest an assumption of parallel trends is invalid.

1.5.2 Attenuation Bias and the Lower Bound

While I believe that the treatment and control groups are more than sufficiently similar to allow for a credible analysis, it is important to recognize that there does exist the potential for bias in the estimates. However, it is quite fortunate that any bias is likely to be only attenuation bias. That is, the findings of a treatment effect are correct, but the true impact may be larger than the estimates suggest. I agree with the perspective of potential attenuation bias to the extent that estimates are likely a lower bound of the true impact. This conclusion stems from 4 fairly simple concerns, that all suggest the possibility of a smaller than actual estimate of the returns to centralized home heating. The primary concerns are: rural home heating in north China, heating above and beyond what is provided by the centralized heating system, general economic wealth of the north versus the south, and outdoor air pollution levels. My first concern is quantifiable and correctable, the second is neither quantifiable nor able to be accounted for given existing data, the third and fourth have been explicitly accounted for in the regression framework

and my findings suggest they are non-problematic.

My first concern is related to the fact that home heating is only centrally provided for cities in the north and not for residents of smaller towns and villages. In the absence of centralized home heating, residents of more rural settings have to resort to traditional home heating, as nearly all Chinese have done until fairly recently. To be explicit, despite centralized home heating in cities, rural residents are still heating using traditional methods and therefore are still likely to be adversely affected by indoor air pollution and low quality home heating. Since the benefits of the central heating system only accrue to urban residents, it must be expected that the benefits would be substantially larger if *all* residents of Northern China, and not merely urbanites, had access to centralized home heating. Stated differently, the estimate is an average treatment effect (LATE) for all residents of North China, however, only urbanites actually receive the treatment (centralized home heating). To get a more refined estimate, we must consider the fact that only 55% of the residents of Northern China are city dwellers, thus it would be appropriate to rescale to obtain a treatment-on-the-treated (TOT) estimate. Doing so suggests the following estimates of the returns to modern heating: Areas with centralized heating have a 0.25% decrease in perinatal deaths, and a decrease of 2.13% in the proportion of low birth weight infants.

Another concern is that heat provision under the centralized heating system may leave some Chinese essentially wanting more heat. This additional demand for heat manifests itself in two ways. One, is that public provision of heat generally only heats buildings to an ambient temperature of roughly 60-62 degrees fahrenheit. It is entirely likely that many residents may desire a warmer temperature, and supplement their home heating with

traditional heating methods. Secondly, additional demand for heat may result from the nature of the "heating season". Recall that centralized heat is only provided during part of the year. Although policy makers adapt the heating season to coincide with the onset of wintry temperatures, there exists the possibility of a lag or simply a period that is cold, but not sufficiently cold to merit public provision of heat. It is quite likely that during these times some households will again choose to supplement with traditional fuels. With either possibility, households are subjecting themselves to additional indoor air pollution. In the absence of this supplementary heating that almost certainly occurs, it would be expected that the estimate of the treatment effect would rise given that northern, urban, Chinese are still being subjected to some indoor air pollution and thus the exposure to the treatment is not entirely complete. This is in contrast to the empirical specification, which supposes that the usage of traditional heating in treatment areas is zero, however there is likely a small, but positive amount of heating in treatment areas that relies on traditional methods. I do not have access to data that allows one to verify, or to quantify the existence and extent of this type of supplementary heating.

Finally, it is important recognize that there are some non-trivial differences between Northern and Southern China that may have a relevant bearing on health matters related to home heating and indoor air pollution. One is that the south is slightly more affluent, and therefore from an economic perspective, better able to finance health care expenditures that may diminish the harmful effects of traditional home heating. This should, to some extent, ameliorate the negative impact of traditional heating. In the absence of this income difference, we would expect that the negative impact would be larger. A simple way

to imagine this would be to imagine a case where southerners, by virtue of greater affluence, all purchase effective home air filtration devices. Findings suggest that these hypothetical air filtration devices could not have eliminated the harmful effects. However, were it not for the air filtration devices we would expect the impact to be larger in magnitude. Of course, greater income need not be spent on air filters, which is merely a convenient example, but rather any normal good that might diminish the negative effects of traditional home heating. Various types of health care consumption come to mind. Despite these concerns, I have explicitly accounted for income differences in the regression framework and it should be noted that any impact attributable to the slight income difference is neither large, nor distinguishable from zero.

Besides a slight difference in overall affluence, there is also a slight difference in the quality of outdoor air. In general, northern China suffers from poorer air conditions than does the south. If northern Chinese are exposed to additional outdoor pollution, this would diminish the magnitude of my estimate of the treatment effect. To illustrate, imagine a case where northern Chinese spend considerable time exposed to outdoor air pollutants. Then we would expect to see some adverse health effects as a consequence, this would thus increase the number of perinatal mortalities and low birth weight infants and diminish my estimate of the treatment effect (since these unpleasant outcomes are occurring in at a greater rate in northern China). However, in no way would this lead to the false conclusion that there is an impact, when in reality the true impact is zero. Again, I have explicitly accounted for outdoor air quality differences in the regression framework, and it should be noted that any impact attributable to the outdoor air quality difference is neither large, nor

distinguishable from zero.

1.5.3 Are Income Effects Present?

Another possible concern is that the provision of centralized heating essentially represented an increase in the real income of northern Chinese urbanites. The potential is that since centralized home heating is heavily subsidized, and therefore quite cheap from the perspective of the consumer, centralized home heating allows for costs savings over previously used traditional heating methods. If households were to take their newly freed income and spend it on health care or other potentially confounding expenditures, it would be possible that it is not really heat provision driving an effect, but rather the increase in real income. Although an entirely plausible possibility, I find little evidence for this explanation. Table 1.2 presents information regarding the costs of traditional and centralized home heating as experienced by households.

The key take aways are two fold. The first key take away is that given the relative prices of traditional methods and modern centralized heating, there is no evidence of systematic costs savings for households. In fact, it appears centralized home heating is slightly more expensive than traditional methods from the perspective of the household. Thus, there appears to be little evidence for heating creating an increase in households' real income; if anything, it appears provision of centralized home heating is associated with a small decrease in real income (roughly $110 at present exchange rates). Moreover, there is very little evidence to suggest that households are anything less than enthusiastic about receiving modern heat given its superior quality of heating, cleanliness, ease of use and freeing up of storage space previously used for fuel. Thus, there appears to be little evidence

for heating creating an increase in households' real income. The second issue worthy of our attention is that due to heavy government subsidization, home heating (regardless of method) is generally a fairly inexpensive proposition that does not consume a large proportion of household expenditures. Given that expenditures are low, we can also conclude that if there were any increase in household income represented by the provision of central heat, it is quite small. It is very unlikely that such a small income shift, if present at all, would be able to generate the types of health effects we see.

1.6 Summary and Conclusion

Few satisfactory measurements have been provided in terms of measuring the social health costs of indoor air pollution. This paper presents evidence of an impact on early-life mortality and birth weight stemming from the smoke produced by traditional indoor indoor heating methods. The primary contribution is the estimation of a credible treatment effect showing the health impact of switching from traditional to modern indoor heating on infants. More generally, this papers provides a small contribution to understanding the social cost of traditional home heating. To do so, I use the central heating system in northern China as a natural experiment, and study the pre-post treatment outcomes within and between groups. The central heating system was introduced in the past 15 years, and only to provinces on the northern side of the *Tsinling-Huaihe* geographic cutoff. Provinces on the northern edge of the cutoff line are the treatment group, while provinces on the southern edge of the cutoff function as the control group. The difference in trends between the two groups is studied using time series of aggregate data from each province.

Using the difference-in-differences approach, I show the existence of a 0.22-0.25 ‰decrease in perinatal mortality for every square meter increase of centrally heated area per person. This translates to about 7,000 fewer perinatal deaths as a direct consequence of centralized heating provision. That is, in the absence of centralized heating it would be expected to have been an additional 7,000 perinatal deaths. As a baseline reference, perinatal mortality is about 0.6% or 6 out of every thousand children die during the perinatal period. I also show the existence of a 0.23%-0.24% decrease in low birth weight infants for every square meter increase in centrally heated area per person. These estimates are significant, consistent and robust to changes in specification. The effects are also plotted in Figure 1.4 and Figure 1.5. Across the treatment and control groups the difference attributable to heating is 1.17%, which translates to roughly 60,000 fewer low birth weight infants in the treatment group than we would otherwise expect in the absence of centralized home heating.

I also closely examine the necessary assumptions that must hold in order for a difference-in-differences methodology to provide credible estimates of a treatment effect. I find no evidence that the parallel trends assumption has been violated. Moreover, I find that trends in heating, perinatal mortality, and low birth weight infants are all quite smooth and thus it is unlikely the treatment effect is generated by random noise in the data. Additionally, I control for potentially time varying differences that may have an impact on perinatal mortality and low birth weights. In particular, I control for GDP growth and outdoor air pollution. Neither proves to be significant. Finally, I consider pre-treatment trends to see if they are indicative of any problems. Due to data constraints, I am relegated

to using a very good proxy in the form of infant mortality. Findings strongly suggest that pre-treatment trends are nearly identical to post-treatment trends, and thus serve to reinforce the central findings.

This paper provides evidence of an effect at the aggregate level. With micro level data sets with birth and death information for infants, it should be possible to identify the seasonality of the effect, as indoor pollution is highly seasonal. Due to data availability constraints, this has to be left for future research agendas. Finally, the elderly may also present a highly relevant and vulnerable population meriting study with regards to indoor air pollution.

Chapter 2

Have Electronic Benefits Cards Improved Food Access for Food Stamp Recipients?

2.1   Introduction

During the past 15-20 years numerous in-kind and cash based government aid programs have moved from paper based vouchers, to electronic ones. Electronic delivery of benefits is predicated on a system where benefits are loaded into an account that is drawn down over a period of time using an electronic benefits transfer card (EBT). Essentially, the system works just like a checking account and a debit card. The key differences being that government agencies provide the funds for the account, while contracted banks provide the EBT card, and manage the account with oversight from government regulators. The most

visible of these reforms has been in the Supplemental Nutrition Assistance Program (SNAP, formerly known as food stamps). In the past, aid recipients would receive monthly booklets of coupons or stamps. These stamps would enable them to purchase approved food from participating vendors at no cost. However, the paper based form of benefits delivery has fallen out of favor for a variety of reasons. Paramount among those reasons was the desire to reduce costs, and streamline the provision of benefits. Although it is clear the switch to EBT cards has been highly successful in reducing costs, there remain questions about other less immediately apparent effects. We intend to determine whether or not the use of EBT cards in the food stamp program, had an impact on the food security of program participants.

Food security is a tremendously important issue from a multitude of perspectives. One distressing aspect of food insecurity is that it afflicts tremendous numbers of children who are susceptible to harm stemming from hunger and compromised nutrition. Murphy et. al (1998) shows that children that suffer from food insecurity are more prone to a wide range of health problems, poor social interactions and have worse educational performance. These problems can result in adverse educational, criminal and labor market outcomes. Paradoxically, Jimenez (2003) and Aliamo et. al. (2001) find evidence food insecurity leads to obesity in many children, and a host of deleterious health outcomes associated with obesity. Thus, food insecurity is clearly not a trivial matter. A commonly held opinion suggests that only the most callous could claim children are to be held accountable for their parents ability to ensure sufficient access to food. Moreover, some argue that in a wealthy society such as our own, food insecurity should be exceedingly rare, rather than a common

occurrence for children who may suffer significant costs and are truly undeserving. This thinking is extended by some to include adults. Although few would argue that everyone is entitled to a certain level of consumption, many have made the case that no one should go hungry in such a prosperous nation. We do not explicitly take a stance on these matters, but clearly hunger is an issue that evokes strong opinions and is an issue few would deem unimportant. Moreover, it is likely a very common problem with the USDA reporting as many as 20% of households received food stamp benefits in 2013. It is within this context that we consider EBT card reforms as having had a potentially costless effect in reducing the specter of hunger. This may be especially important given that the USDA reports 76% of food stamp households contain children.

From a theoretical standpoint it is not immediately clear whether or not a change in the method of delivery benefits would have an impact on food insecurity. On the one hand, benefit levels do not change, food prices are independent of benefits deliveries, and there are no surcharges or transaction costs associated with use of an EBT card. However, there are two reasons to think that the transition to food stamp benefits being delivered via EBT might have lasting impacts on food insecurity.

One consideration is that paper based food stamps were relatively more prone to fraudulent sales of benefits. If program participants sell fewer benefits, they would likely use more benefits for the intended purpose of food purchases. This would likely lead to a reduction in food insecurity that would persist for as long as EBT cards diminished food stamp trafficking. The second primary channel would be that many food stamp program participants may suffer from loss and theft of benefits. Under a paper based regime, a

victim of loss or theft would have to wait until the next month to receive new benefits. Under an EBT card regime, the card could be replaced in short order and the benefits would not be able to be stolen without knowledge of a secret PIN number. Together EBT cards should would make loss less costly, and theft less likely as well as less costly. Finally, a transitory impact might be possible from the standpoint that program participants may not immediately know how to make a painless transition from paper based benefits to an EBT card. Some recipients may fail to receive or collect the card in the first months, they may struggle to use the card and they may fail to learn where it can be used. These considerations would likely lead to an increase in food insecurity, but any increase would be merely transitory as program participants would either acquire their cards and/or learn how and where to use the card.

We consider the roll-out of EBT card reforms across California counties during the period from 1997 to 2006. We link data detailing the county roll out of EBT card reforms with self-reported food security measures using the California Health Interview Survey. Furthermore we use a variety of data sources to control for program participant characteristics, county characteristics, benefit levels and economic conditions. Our identification rests on the quasi-random timing of EBT card reforms at the county level. Our findings are generally not supportive of the claim that EBT reforms altered food security for Food Stamp Program participants. We do however find substantial evidence of a very short lived increase in food insecurity at the moment of EBT card reform implementation. In particular, food insecurity becomes roughly 2% more likely for our preferred measure of food insecurity. Effects last only one or two months before disappearing. This transitory

effect suggests individuals may have had an initially difficult transition to the new system that resulted in an increase of food insecurity lasting between 1 and 2 months. This effect is robust to a variety of specifications and measures of food insecurity. However, as stated above, there is no systematic evidence that EBT cards have altered long-run food security. The paper is organized as follows: section 2 considers relevant academic literature, section 3 contains the necessary background information for understanding EBT card reforms and California specific details of the Food Stamp Program, section 4 presents and discusses the data sources used, section 5 offers analysis and discussion of our results, section 6 offers some concluding remarks.

2.2   Literature Review

Although a considerable literature with regards to food stamps exists, and there are papers that considers the costs and benefits of EBT cards, very little to our knowledge has been written investigating a specific linkage between EBT cards and food security. More generally, there has been relatively little work done looking at externalities, or unintended effects of EBT reforms. Despite this general void, two recent papers have considered the possibility of a relationship between EBT cards and crime. Lovett (2014) and Wright et. al. (2014) use the roll out of EBT card reforms in California and Missouri, respectively to determine if EBT cards had any bearing on criminal acts. Lovett considers the effect of lost cash income on crime. The income loss stems from decreased fraud opportunities as a result of EBT card reforms. Wright et. al. consider EBT card reforms as decreasing the cash holdings of food stamp households through the same channel - diminished sales of

food stamp benefits. Lovett reports an increase in some crime categories, while Wright et. al. report an overall decline in crime.

Apart from these two papers linking EBT to crime, there has not been an explicit research agenda investigating externalities associated with EBT card reforms. However, many papers have made key contributions to the current understanding of the Food Stamp Program, and EBT card reforms.

2.2.1 Paper-based Delivery versus Electronic Delivery

An older literature, mostly from the 1990's, considers the benefits and costs of using EBT systems to deliver benefits. In general, the literature has been very supportive of the use of EBT cards and it is quite probable that the benefits resulting from EBT cards substantially contributed to the decision to mandate EBT cards as part of the Welfare Reform Act of 1996. Some of the benefits noted by Fraker et. al (1992) and Ohls et. al. (1992) are that EBT cards have been a success in terms of cost reduction, better record keeping, and provision of recipient services.

Andrews and Wilde (2000) offer a meta-analysis of much of the work done in the 1990's. A central finding that is pertinent to this paper is that they find EBT offers significant benefits from an anti-trafficking perspective. This stems from the consequence that with EBT cards, it is much easier for the USDA to identify suspicious patterns in benefits redemption. A common red-flag is if a store that sells relatively few groceries (and/or at comparatively high prices) has disproportionately large purchases occurring at the beginning of the month. If this is a common occurrence the store is likely to be investigated. Should USDA investigators conclude that the store engaged in trafficking, the store will

likely lose it's ability to accept food stamps. Furthermore, investigators are often able to observe which households are engaging in suspicious behavior and pursue action against those households.

Andrews and Wilde (1994) also find evidence that there are households that choose to engage in trafficking even though they often run out of food before the end of the month. Moreover, they find that within months the pattern of benefits usage is often far from smooth. They note that the dominant trend is that benefits are used in large quantities at the beginning and end of the month. In particular, the largest day for benefits use is the 27th day of the month. They document that many households cautiously allocate their spending throughout the month and upon learning there are ample benefits to last until the next release of benefits, they spend the bulk of their balance. This typically occurs near the end of the month. Although this behavior typifies many households, another important subset is likely to use the bulk of their benefits at the beginning of the month. A surge of benefits usage at the beginning of the month is consistent with having run low on, or out of, benefits the previous month. Having run out is by no means a sufficient condition for having trafficked, but is often very much related. Consistent with this interpretation is that bulk of end of month purchases occur in supermarkets (locations that generally are far less likely to traffic benefits, while a much smaller share of the beginning of month surge occurs in supermarkets, with many more transactions occurring in small convenience-type stores (which are generally far more likely to traffic benefits).

Macaluso (2000) reports that EBT was generally much favored by program administrators as a result of cost savings and general ease of use for program participants.

49

A USDA report (1994) states that EBT was also popular with recipients as they reported diminished shame costs and lowered risks associated with loss or theft. Danielson and Klerman (2006) make use of administrative data, and find that EBT, paired with a simplified certification process, led to a small increase in program participation. Hanratty (2006) was able to replicate this finding and found EBT card reforms coupled with eased re-certification increased program participation by about 6%. It should be noted that for both of these studies the authors are unable to separate effects resulting from EBT and altered certification. Interestingly, Haider et. al. (2003) find that the elderly were generally unaffected by these changes and did not participate in the program at altered rates.

2.2.2 Cash versus In-Kind Transfers

At one time a very active research agenda related to whether it was a good idea to have the Food Stamp Program as an in kind transfer. As a result there are numerous papers considering the likely impacts of changing food stamps into a cash transfer program. This is often referred to as "cash-out". Food stamps has always been an in-kind transfer for several reasons, but the most important has been the desire to ensure benefits are spent on actual food and not purchases of other goods and services such as alcohol or tobacco. This is especially true in households with children, which constitute the bulk of food stamp households. Of course, it is not clear that the goods and services bought would necessarily be alcohol or other potentially problematic goods. The arguments for cash out are that food stamps are more costly to administer as an in-kind transfer, and food stamps are often distortionary. In this case distortionary means that recipients consume more food than they would under a cash transfer. This is a matter of some importance because distortions may

create a considerable incentive to engage in food stamp trafficking, which is potentially a key channel for impacting food security.

This question of distortion has primarily been addressed using observational studies, summarized by Fraker (1990), in which researchers estimate the marginal propensity to consume food using the following specification:

$$food\_spending_i = \beta_0 + \beta_1 Cash_i + \beta_2 fstamp_i + \mathbf{X}_i \gamma + \varepsilon_i$$

Here the unit of observation i is the household. $Cash_i$ and $fstamp_i$ are income from cash and food stamps, respectively; while $\mathbf{X}_i$ is a set of household observable characteristics, and ε_i is the error term. Under this framework, the differential impact of food stamps as an in-kind transfer is the difference between the estimates of β_1 and β_2. The difference between the two coefficient estimates reveals the extent to which food stamps create extra food consumption that would not have occurred under a cash transfer. These studies offered important and interesting conclusions, but are currently viewed as suffering from general problems associated with observational studies that do not randomize treatment exposure. In particular, program participation is taken as exogenous. Related to the issue of exogeneity both Currie (2006) and Moffit (1983) study models of program participation and find that participation is effectively a choice variable. Thus, individuals choose to participate or not and that choice is correlated with preferences for food consumption over other forms of consumption. Moreover, comparisons are made between eligible families who participate and those that do not participate. These problems have led to the present belief that estimates of distortion suffer from bias.

In addition to observational studies, there are two papers investigating the impact of pilot programs (in Alabama and San Diego) initiated by the USDA on the consequences and feasibility of cash out. Fraker et. al. (1992) and Ohls et.al. (1992) report estimates of distortions being about 5%, meaning that households receiving in-kind benefits had 5% higher expenditures on food than households that received a cash transfer.

A more recent trend is exemplified by papers written by Hoynes and Schanzenbach (2007 and 2009) that rely on the initial roll out of the Food Stamp Program. It should be noted that the setting is meaningfully different than the randomized experiments in Alabama and San Diego, in that the initial launching of the Food Stamp Program resulted in households receiving a real income shock. In comparison, the randomized experiments of the early 90's held income constant. Only the form of benefits delivery is altered, not household purchasing power. The central findings of papers using the historic roll-out of the program, is that stamps or cash do not constitute a meaningful difference for infra-marginal participants (households that spend some cash on food). However, cash transfers will diminish food expenditures for constrained or extra-marginal households (households that only use food stamps to purchase food). Thus, distortion occurs for extra-marginal households, but not infra-marginal ones.

In terms of quantifying this split in household types, Whitmore (2002) finds that roughly 20% to 30% households are infra-marginal. These infra-marginal households would be relatively unaffected by a transition to a cash only system of benefits. Whitmore also uses survey data to suggest that a cash transfer may result in less aggregate spending on food, but that most of the reductions occur on beverage spending. The implication being

that nutrition is unaffected. Although it is not clear that nutrition is unaffected, calories ingested do decrease.

Whitmore notes that for program participants with a distorted preference for food (as a result of the in-kind nature of the transfer) value their benefits at roughly 80 cents on the dollar. This valuation is important as it implies that roughly one fifth to one third of program participants would potentially have an interest in trafficking. A potentially problematic issue is that the study doesn't make use of pre-intervention data and as a result the definition of an infra-marginal household is a household that doesn't use all its food stamps. Very basic consumer theory strongly implies this should never occur, unless households sell some of their benefits. The paper does not offer a choke price for extra-marginal households. This price would presumably be lower than for infra-marginal households, but to what extent is unclear.

Butler and Raymond (1996) also consider the issue of nutrition, and whether or not preferences for food consumption cause some individuals to select into the Food Stamp Program. Using a sample of seniors within a cash out pilot program, they find that compared to cash, in-kind benefits do not offer any advantages. In comparison, Bastiosis et. al. (1998) note that the Food Stamp Program is quite beneficial from a nutrition perspective, but findings rely on older data from the 1980's and they do not consider the cash versus in-kind question.

Blundell and Pistaferri (2003) consider the possibility that food stamp income may be useful as a consumption smoothing mechanism. They find that food stamps do aid households in smoothing consumption in the face of income shocks. However, following

a large negative income shock they find that households using food stamps consume 30%
more food than they would had they received cash instead. Although this large negative
income shock may be viewed as something of a special circumstance, this implies a higher
degree of distortion than in other papers.

Although it may not be immediately apparent, the findings of the cash or in-kind
debate is relevant for us in that if many recipients spend some of their income on food, there
is a much diminished incentive to traffic food stamps. This results from the simple fact that
if households engage in trafficking, they must always sell their benefits at some discount. If
program participants are spending some of their own cash on food, selling benefits would
result in them needing to buy the same amount of food, but with less purchasing power to
do so, since they have sold benefits at less than full face value. As a result, any cash gains
would be more than offset by the loss of purchasing power. Thus, trafficking would be a
rational option only for constrained (extra-marginal) households that never buy food with
their own cash. Extra-marginal households may find that trading some portion of their
food benefits for greater consumption flexibility is warranted.

It should be noted that some feel the above thinking is too simplistic in that it
ignores the within month variation in food stamp benefits and cash availability. Essentially,
at the beginning of the month households may have a relatively high ratio of food stamp
benefits to cash. A household that heavily discounts the future, may choose to sell some
benefits for cash. However, later in the month after food stamp benefits are exhausted,
the household may use some cash to purchase food. In this case a household that spends
some of its cash food, may still choose to traffic in food stamps as a result of the changing

marginal utilities of cash versus food during the course of the month.

2.2.3   How Much Trafficking Occurs

A natural question for us is: "To what extent are food stamp benefits actually trafficked". This is important since a key channel for altering food insecurity are changes in the extent of food stamp sales. Households that sell some of their benefits are clearly more likely to experience food insecurity than if they had not sold some of their benefits. If EBT card reforms diminish sales of benefits, this should translate to a reduction in food insecurity. Thus, the frequency with which trafficking occurs will likely be important in determining the magnitude of any food insecurity decreases. It is worth noting that the only case in which food insecurity would not fall, is if only households with ample food resources engaged in trafficking prior to EBT card reforms. If households with ample food stuffs are the only households trafficking benefits, then a reduction in fraudulent sales may not result in improved food security. This can also be thought of from the perspective of distortion. Distorted households will be more likely to traffick and EBT card reforms may make them even more distorted. These distorted households are less likely to experience food insecurity.

Anecdotal evidence and stories regarding food stamp misuse, fraud and trafficking abound and despite the lack of rigor, policy makers are very much interested in knowing how this market functions, and to what extent it exists. Whitmore (2002) uses survey data to suggest that food stamps are valued at roughly 65 cents on the dollar. This is an interesting implication in that 65 cents is not sufficient to induce the average infra-marginal household to sell benefits, which she estimates values its benefits at 80 cents on the dollar in

previous work. However, since we don't have an idea what the valuation of extra-marginal households is the estimate of 65 cents does not tell us whether extra-marginal households would likely be interested in trafficking or not. Moreover, the estimate does not account for the possibility that the value may change throughout the course of a month.

Macaluso (2000) investigates data from more than 10,000 USDA investigations of stores suspected of black market activity. Findings claim that roughly 4% of all benefits issued can be traced to misuse in investigated stores. The USDA's current position is that EBT cards have reduced this number, but it is not known to what extent this has occurred. Macaluso also notes that smaller stores are substantially more likely to engage in trafficking than supermarkets. Finally, the suspected degree of trafficking fell during the 1990's with the author attributing much of this decline to EBT cards, which made large scale trafficking much less likely. This last assertion is important with regard to this paper as it supplies empirical evidence for the USDA's claim that EBT cards increase the risk of a food stamp trafficker being caught and therefore would push down the price for black market food stamp sellers. A key point to keep in mind is that Macaluso offers an estimate of the rate of trafficking that involves stores and does not account for person to person sales. Since this avenue is not considered, 4% might safely be thought of as a lower bound.

An interesting aside worth considering is that if EBT reduces trafficking and in doing so reduces food insecurity, this may not always be an unambiguously desirable outcome. Although we imagine food security as being a paramount concern, there exists the possibility that cash generated by trafficking is used for other goods and services that could be potentially quite important. Examples might be households that sell benefits and suffer

some food insecurity, but in doing so generate cash income used for medication or utility services. Given the nature of the Food Stamp Program as an in-kind transfer, a consequence is that distorted households will consume more food than if they had received cash. These households presumably value other goods and services more highly than food. There exists the distinct possibility that even from a paternalistic perspective, we may not always view higher valuation of other goods and services as problematic. Clearly, if a household chooses to sell benefits for sin goods we would likely view this as undesirable. However if a household sells benefits to finance the purchase of necessary medication, the normative conclusion is far less clear.

Ciemnecki et. al. (1998) addresses the question "Is it possible to rely on survey data to obtain a measure of food stamp trafficking?" The authors rely on several different surveying techniques, and also make use of ethnographic methods, to ask food stamp recipients about their behaviors and attitudes regarding fraud and trafficking. It should be noted that although the authors feel that none of the individual methods were entirely conclusive, all of the methods employed suggest that individuals considerably under-report trafficking. In terms of important take aways, we note that self reported trafficking was lower in areas that used EBT cards than in areas that used paper stamps. An additional finding was that where EBT cards were in place, respondents frequently stated that this made selling benefits to other individuals much more difficult, as it required the buyer and seller to redeem the benefits together.

2.3 Background

2.3.1 Historical

In 1961 President Kennedy announced the creation of a pilot food stamp program in eight counties and the modern Food Stamp Program (FSP) was created. In the following two years the pilot programs were expanded to 43 counties and successes with these pilot programs led to the Food Stamp Act of 1964. The Food Stamp Act gave county governments the authority to start up the FSP in their county. As is still the case today, the program was federally funded and benefits were redeemable at approved stores. With considerable press attention, demand for food stamp programs grew and this interest culminated in passage of the 1973 Amendments to the Food Stamp Act, which required all counties to offer food stamp benefits by 1975.

The food stamp program provides qualifying individuals and households with vouchers to be spent on approved food items with participating vendors. Since 2008 the FSP has been known as the Supplemental Nutrition Assistance Program (SNAP). As of 2013, the average participating household receives slightly more than $200 a month in benefits. With regards to program administration, states have considerable freedom to structure the program as they see fit; thus there are considerable differences in implementation across states. However, so long as states meet federal standards the final source of funding is at the federal level. As a result it is not entirely correct to think of the program as purely a state or federal program. California is different from the other 49 states in that its food stamp program is administered at the county level, as opposed to the state level. This means that county Social Services offices are responsible for determining eligibility, disbursing benefits

and assisting beneficiaries with program related difficulties. California is still accountable to Federal requirements and is still federally funded, but the day-to-day operations are conducted at the county level, whereas elsewhere the program is administered at the state level.

2.3.2 Reforms

There have been considerable changes and reforms to the FSP during the past 15-20 years. This paper aims to investigate food security changes stemming from changes in the delivery of benefits. Until the late 1990's food stamps consisted of small rectangular pieces of paper with dollar value denominations. A recipient would receive a booklet of stamps at the beginning of the month that could be used in place of cash to make approved food purchases (see Figure 2.1). However, today no one uses actual paper stamps. In their place, benefits delivery has moved to an electronic system where beneficiaries receive an electronic benefit transfer (EBT) card linked to an account that contains funds to be used for appropriate purchases (see Figure 2.2). In essence, the system functions like a checking account and a debit card. At the beginning of each month benefits are loaded into the account and the individual is free to use those funds to make purchases.

The first EBT trial was in Reading, Pennsylvania in 1984. Starting in the early years of the 1990's several pilot programs were conducted to evaluate the feasibility of EBT benefits delivery. By the end of the decade EBT cards had been adopted in several states. In 1996, as part of the Personal Responsibility and Work Reconciliation Act (generally known as Welfare Reform), the Federal government required all states switch to use EBT cards by October 2002. Compliance was not achieved until the end of 2004. Although there

were threats of withheld funds, in practice this did not occur and there were no disruptions in funding to states that converted after the October 2002 deadline.

2.3.3 Reasons for Reform

The shift from paper stamps to an electronic system was motivated by a variety of concerns. Most prominent among many considerations was cost as paper food stamps were rather costly from an administrative perspective. Among the most interesting of these costs were measures taken to diminish the ease and impact of counterfeit stamps. Examples of these costly measures were intaglio printing and fairly sophisticated water marks (see Figure 2.1). In addition to anti-counterfeiting measures, allocation and distribution of the stamps were fairly costly. Some states made it obligatory for recipients to come into the local social services agency and receive their stamps in person. This imposed considerable costs on program participants. In other states the preferred method was delivery by mail. This had its own problems in that it was vulnerable to theft and also created non-trivial mail delivery costs. Mail delivery was also problematic for individuals without an address or a stable one. Another issue was that because food stamps were basically as easy to use as cash, mailing food stamps was not terribly different from mailing cash in the sense that there were considerable opportunities for theft. By switching to EBT cards, Social Service agencies would also be able to avoid imposing loss and theft associated with mail based delivery on their clients. Similarly, repeated opportunity costs associated with requiring that their clients come to the social service agency to pick-up stamps in person could be avoided.

For program participants, an important consideration was that paper-based stamps

Figure 2.1: Food Stamp

Figure 2.2: New York EBT Card

were vulnerable to loss and theft. In the case of loss, the month's supply of food stamps were just that – lost. With an EBT card a beneficiary that loses their card has not lost the funds associated with the card (provided no one has used the lost card, which requires knowledge of a PIN number). They can request a new card be issued to them, and the new card will still have the same account balance as the old card. Along similar lines of thinking, theft is a considerably diminished possibility and a potentially less costly one. Under the old regime of paper-based stamps, stolen stamps could be used by just about anyone. However, with an EBT card, a thief would not be able to debit the account, despite physical possession of the card unless they also know the PIN number. In some cases cards showed a photograph of the correct owner, making stolen cards even harder to use (see Figure 2.2). These questions of loss and theft are quite important from a food security perspective in that someone who has lost, or had their benefits stolen, would be much worse off under a paper based regime than an EBT card regime. They would potentially be unaffected under an EBT card regime as the card can be replaced quickly and nearly costlessly. Theft of a card would be virtually meaningless, unless the thief also knows the PIN number.

Fraud was another motivating factor for the move to an EBT system. Since stamps could be used by anyone, there was considerable ease in re-selling stamps. Of course

most beneficiaries were content to use their benefits as intended (for the purchase of food). However, many might have a marked preference for the greater consumption flexibility afforded by cash. These individuals would likely be willing to sell their benefits at less than full face value in exchange for cash. It is for this reason that the food stamp program was specifically structured as an in-kind transfer, and not a cash transfer. That is the paternalistic desire to bypass some of these recipient preferences for non-food consumption. Clearly EBT does not represent a total solution, but one that diminishes the ease and attractiveness of selling benefits. With an EBT card, the seller of benefits would have to accompany the buyer during their purchases or sufficiently trust the buyer to give them the PIN. These additional complications and inconveniences would presumably lower the value of food stamp benefits and decrease the number of transactions.

It is worth noting that the most common form of food stamp trafficking is not the person to person example described above. Instead the program beneficiary exchanges benefits for cash (or unapproved goods) with an approved retailer of food. In the past, the transaction would be as simple as a recipient swapping stamps for cash. The grocer then would redeem the stamps with state authorities for the cash value of the food that was purportedly sold, but in reality was not. However, this type of transaction becomes riskier under an EBT card system. The simple reason for this is that since EBT is electronic and the sales data is retained by the state, any irregularities or suspicious aspects of the sales data could result in investigation of the suspected grocer. To quote the USDA on the matter:

> SNAP electronic benefits transfer (EBT) has given USDA new tools to iden-
> tify, track, and take action against trafficking. We use the electronic "audit
> trail" from EBT transactions to identify trafficking and other suspicious activ-

ity. The Anti-Fraud Locator using EBT Retailer Transactions (ALERT) system monitors electronic transaction activity and identifies suspicious stores for analysis and investigation. FNS (Food and Nutrition Services) has a dedicated team of over 100 analysts and investigators across the country dedicated to SNAP retailer compliance. They analyze retailer data, conduct undercover investigations, and process cases – including fines and administrative disqualifications – against violating retailers.

The issue is potentially very relevant from a food security perspective as food stamps trafficked are likely used for non-food consumption. As such, food stamp trafficking may potentially increase food insecurity for individuals choosing to sell benefits for cash. By diminishing fraud, EBT may result in households having greater real income for food purchases, thus diminishing food insecurity.

It is worth noting that directly altering food insecurity was not a primary motivator for the transition to EBT cards. There was, however, a secondary channel considered that may have had an indirect impact of food security. Policy makers and advocacy groups have been long concerned with what are generally termed shame costs. These shame costs may prevent many eligible program participants from applying for and receiving food stamps. The idea behind shame costs is simply that using food stamp benefits is a publicly observable activity that can be a source of shame for many people. For some marginally inclined people, these shame costs may be sufficient to alter their binary decision to apply for and use food stamp benefits. Use of an EBT card is far less obvious than the use of stamps. Thus, any shame would be greatly reduced as others would not be privy to knowledge about food stamp usage. If reduced shame costs mean more program participants, this would likely reduce food insecurity in the general population. Thus, EBT cards may create a selection effect that could potentially reduce food insecurity.

A final consideration motivating EBT card reforms was the desire to streamline the provision of benefits to individuals receiving aid under multiple social welfare programs. A sizable subset of food stamp recipients also receive assistance from other programs such as general assistance, TANF (Temporary Assistance for Needy Families) and SSI (Supplemental Security Income). Many states have chosen to link these accounts together with a single EBT card.

2.3.4 California Specifics

In terms of the food stamp program, California is unique from the rest of the nation in one key way. California administers its FSP (known as Cal-Fresh within California) at the county level, rather than at the state level. As a consequence, when the State Department of Social Services mandated transition to an EBT system, it was left to each individual county to reach compliance. This resulted in considerable variation in the timing of reform implementation, with some counties adopting EBT years before other counties. This variation in the timing sets the stage for a potentially attractive natural experiment.

However, such a strategy would be severely undermined if the timing of the reform was endogenously determined by factors that correlate with food insecurity outcomes. Table 2.1 offers evidence that this was not the case. Despite this evidence, the causes of variation in timing of reform are of considerable interest.

Table 2.1: Early and Late Reforming County Comparison

	2000 Decennial Census				2005 American Community Survey			
	Group 1	Group 2	Group 3	Group 4	Group 1	Group 2	Group 3	Group 4
Mean Age	33.64	34.52	35.61	38.13	35.94	36.94	38	39.64
	(21.92)	(22.13)	(22.64)	(22.91)	(22.35)	(22.75)	(23.02)	(23.16)
Mean Family Size	3.49	3.56	3.26	3.14	3.45	3.42	3.19	3.15
	(2.10)	(2.14)	(1.98)	(1.72)	(1.98)	(1.90)	(1.80)	(1.60)
Median Income	27000	30000	27000	32000	33200	35000	32000	39000
Food Stamp Use Rate					0.049	0.051	0.081	0.013
Percent White	0.629	0.61	0.7	0.86	0.671	0.645	0.72	0.845
Percent Black	0.064	0.046	0.069	0.01	0.053	0.038	0.057	0.008
Percent Asian	0.076	0.126	0.086	0.03	0.096	0.147	0.101	0.051
Percent Hispanic	0.08	0.298	0.149	0.14	0.317	0.278	0.155	0.129
Percent Married	0.401	0.416	0.402	0.465	0.424	0.435	0.431	0.501
Percent Employed	0.439	0.442	0.426	0.483	0.45	0.447	0.443	0.465
Urbanicity Measure	54	61	49	31	54	64	48	33
No High School	0.394	0.402	0.372	0.331	0.36	0.357	0.328	0.299
Percent Citizen	0.878	0.847	0.92	0.928	0.886	0.874	0.928	0.942

County groupings determined by timing of reform. Group 1 earliest counties with each successive group reforming later. Group 4 being the latest.

Group 1: San Diego and San Bernardino counties.

Group 2: Alameda, Contra Costa, Fresno, Humboldt, Imperial, Kern, Marin,Merced, Orange, Riverside, San Francisco, San Joaquin, San Luis Obispo, San Mateo, Santa Barbara, Santa Clara, Santa Cruz, Solano, Sonoma, Stanislaus,Tulare, and Yolo counties.

Group 3: Butte, Colusa, Del Norte, El Dorado, Glenn, Inyo, Kings, Madera, Mono, Monterey, Nevada, San Benito, Shasta, Siskiyou, Sutter, Tehama, Trinity, and Yuba counties.

Group 4: Amador, Alpine, Calaveras, Lake, Mariposa, Modoc, Napa, Placer, Plumas, Sierra, and Tuolumne counties.

2.3.5 Sources of County Timing Variation

One issue of substantial consequence on the timing of reforms was the matter of whether to use EBT for food stamps alone, or to use EBT with other governmental aid programs as well. The Welfare Reform Act of 1996 stipulated that food stamps be delivered via EBT, but there were no such requirements that other programs use EBT. Despite the lack of a federal or state mandate slightly more than half of California counties chose to deliver benefits from other programs via EBT as well (37 of 58 counties, generally more populous). Converting other aid programs likely made reform implementation more complicated and increased the time needed to convert to EBT.

Another consideration was what is known as the food stamp stagger and differences across counties in use of the food stamp stagger. The stagger essentially works in the following way. Food stamps benefits are usually not all released on the 1st of the month to all beneficiaries. Instead some beneficiaries receive their benefits on the 1st of the month, while others receive it on the 2nd and so on until the last recipients receive their benefits on the last day of the stagger (usually the 10th day, but as late as the 25th in some counties). Different counties used different algorithms for deciding which recipients would receive their benefits on which days. Once the state mandated EBT, they also required that all counties use an algorithm based on the last digit of the case number. In many counties, this was quite problematic because switching the algorithm resulted in some people going more than 31 days without new benefits. This is prohibited by Federal Law. In these instances agencies had to issue emergency food stamps (paper based), and more generally, prepare for a rockier transition to EBT.

The issue of mail-based delivery or over-the-counter issuance also contributed to timing variation. When issuing the first EBT cards, counties had to choose to mail out cards and PIN numbers, or instead to have recipients come into County Welfare Department offices to receive cards and PIN numbers (as well as training in some cases). This decision had to be reached independently for both the newly entering participants, and those that were continuing in the FSP. Different counties opted for different methods, with Over-the-Counter issuance being generally much slower.

Language support services were another factor creating variation in reform timing. California as a state provided support materials for program participants in 10 languages. Despite this some counties felt that they had significant populations that would need assistance in other languages. These counties took it upon themselves to translate documents and brochures into additional languages. This additional work likely delayed conversion to EBT cards.

An imperative fact to bear in mind is that EBT card services were contracted out to private banks, not directly provided by the state. Counties had to reach an agreement with a private bank. Given the complex nature of the task, these negotiations often took a great deal of time. Some of the sticking points in negotiations between banks and county officials were: surcharges for cash benefits access, transaction fees for cash benefits access, cash benefits stagger (3 day versus not at all), mailing based issuance of EBT cards and associated costs, customer/recipient support services, availability of acceptable ATMs for Limited English Proficient recipients, loss/theft replacement services, and language issues (i.e. what languages were to be included in mailings and phone support lines, both

automated and personal).

2.4 Data

We rely on two essential sources of data in order to determine if EBT card reforms have had a bearing on food insecurity for food stamp beneficiaries. One data set, provides us with information about California's roll out of EBT card reforms, while the other allows us to contemporaneously track food insecurity across California. The central condition for both data sets is that they need to have time information at the monthly level and they need to have geographic detail down to the county level. Given that this information is available, the two data sets can be linked to answer questions about changes to food security at the time of EBT card reforms.

2.4.1 CALFresh Data

Food stamp reform data was obtained via a unique data set that was obtained by conducting telephone interviews with each county Social Service agency in California. We called each county agency in California and conducted a phone interview with county officials to determine the timing of reforms, as well as a host of relevant contextual information. That contextual information included: whether or not the county used EBT for cash aid programs, whether or not the county converted under state order or prior to the state order, whether the county reached compliance in time or not, if EBT cards where issued in person, by mail, or both, if a change in delivery methods occurred at the time of EBT reforms, and if any other changes to the programs administration coincided with EBT card

reforms. Finally, we asked the county official if they believed there was anything unique to their county's conversion process. This generally did not elicit a response, but in some cases it was revealed what the particular difficulties were in converting to an EBT card regime. However, the most important piece of information was the month and the year in which the county converted. This is vital as it allows us to determine if food insecurity occurred in a paper based food stamp regime or an EBT regime - in effect whether a county at a given time belongs in the treatment group or the control group. In some cases the title of the agency varies. For example, some counties implement food stamps through the office of Health and Human Services, while others have the title of Social Services. Nonetheless, the task of the county agency remains the same. Along the same vein, interviews were generally conducted with an official having the title of " CAL-Fresh Program Manager" or "CAL-Fresh Program Coordinator". Thus, the title varied, but the nature of the job remains the same.

2.4.2   California Health Interview Survey

The second data source used is the California Health Interview Survey or CHIS. The CHIS is modeled on the well-known and oft used National Health Interview Survey, but is intended to give greater detail and within the confines of California. The CHIS is a random state-wide health survey conducted every two years since 2001. Participants in the CHIS survey are chosen at random, and the sample is extensive enough to be statistically representative of California's diverse population. CHIS telephone surveys are conducted in all 58 counties of California and in person interviews have been conducted in all counties, but not in every biennial wave.

CHIS covers a great many health issues, but importantly for us asks a series of questions pertaining to food security. Moreover, these are considered core questions, and as such are repeated in each survey in order to measure significant shifts over time. These food security questions ask about the frequency in which the respondent has run out of food due to lack of money, couldn't eat balanced meals due to lack of money, eaten less than they should, or skipped eating some meals to ensure food would last throughout the month or food would be available for dependents. This is highly valuable as it allows us to consider a wide range of food security measures that could be potentially impacted by changes to the food stamp recipient population.

CHIS also allows us to observe a great deal of individual level detail about the respondent such as education, age, income, gender, food stamp program participation, WIC program participation, known health problems, as well as racial, demographic and some economic information. Thus, we are able to explicitly control for a host of potentially confounding attributes in our analysis using CHIS data.

CHIS data does have some notable shortcomings for the purpose of our analysis. Notably CHIS data begins in 2001, thus we have post EBT reform information for all 58 counties, but we only have pre-reform data for 54 counties. San Diego, San Bernardino, Alameda and Yolo counties are the adversely affected counties. According to USDA definitions, San Diego, San Bernardino and Alameda are all "large counties by food stamp program population". Another shortcoming is that although CHIS is representative at a state level, it may have underreporting for some smaller, low population counties. Thus, there may be some bias regarding estimates related to smaller counties.

2.4.3 Other Data Sources

We incorporate information regarding the changing nature of the food stamp program using USDA SNAP data relating benefits levels and participation levels within California over time. The USDA data allows us to explicitly control for differing benefit levels across counties and over time. We use the decennial Census and ACS to account for changing economic and demographic conditions at the county level.

2.5 Analysis and Results

2.5.1 Necessary Assumptions and Conditions

Before considering regression analysis it is important to verify that the empirical environment is indeed appropriate for such techniques to be used. The underlying assumption that allows for our analysis is that the timing of the reform is largely exogenous to within county food insecurity. If it is the case that some counties reform earlier because the county's population is disproportionately predisposed to food insecurity, and county administrators predicate their conversion timing on this information, we are no longer able to say that the timing of the reform is exogenous. Without exogenous timing of the reform, we lose the quasi-random timing assumption and our estimates will suffer from bias. To be explicitly clear from an empirical standpoint, counties belong to the control group so long as they are using paper based benefits delivery, and move to the treatment group once EBT card reforms are implemented. The empirical strategy would suffer if our control groups did not serve as credible counterfactuals for the changes in the treatment group. Thus, it cannot be the case that early reforming counties are radically different than late reform-

71

ing counties, and conversely late converting counties should not be radically different than early converting counties. Since each group serves as a counterfactual control group it is necessary that they essentially be good controls.

Table 2.1 presents evidence that timing of the reform is indeed exogenous to relevant observable characteristics. We divide counties into four groups based on the timing of the conversion from a paper-based system to an EBT system. Figure 2.3 reveals the geographic distribution of these groupings. Group 1 is composed of the earliest reformers and subsequent groups converted at a later point in time, with group 4 being the group of counties that reformed the latest. Table 2.1 shows the means of various relevant observable characteristics such as food stamp usage rates, crime rates, unemployment rates, demographic characteristics, and other potential pertinent characteristics. The central finding is that counties across groups are not systematically different across observable characteristics barring urbanicity. In general, early adopters tend to be larger and more urban counties than late adopters. The potentially reflects the fact that larger county welfare agencies likely have more resources that they are able to bring to bear, and thus were generally able to convert more quickly. There are, of course, notable exceptions. San Bernardino County (an extremely sparsely populated county) was among the first to convert, while Ventura County (a quite urban county) was among the late reformers. An important take away is that it is very unlikely that the timing of EBT card reforms are endogenously determined by the severity of food insecurity problems.

Figure 2.3: Map of EBT Card Conversion Timeline
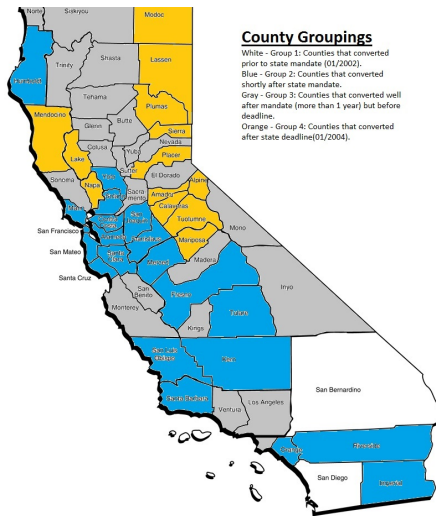
2.5.2 Estimation

We use a difference-in-difference (DID) framework to estimate the impact of EBT card reforms on food insecurity. Our framework is slightly more general than the two period form of DID estimation. This is largely due to the more complex nature of the rollout of EBT card reforms across California. The circumstances are different in that rather than having two groups in two time periods, we have 58 counties that all change from being paper based counties (control counties), to being EBT counties (treatment counties), and the transition from control to treatment group is specific to each county and can occur in any month from January 1998 to December 2006. Given the nature of spatial and time variation, we use a form of the DID technique commonly referred to as an event study framework where we consider changes to food security in each county as it transitions from a paper based regime to an EBT card regime. This greater degree of complexity suggests a more flexible specification may be beneficial. Presumably an event study framework allows

for considerably more nuanced results since it more closely accounts for the fact that we effectively have 58 quasi-experiments occurring at different points in time. Finally, and perhaps most importantly, we prefer an event study framework because it allows us to consider the treatment effect as dynamic over time and not static, as is the case with a two period DID framework.

We consider a variety of estimation techniques including ordinary least squares as a linear probability model (LPM), probit and logit. Estimates do not vary substantially across specifications, however we follow Horace and Oaxaca (2006) in preferring the probit specification as it is better able to handle measurement error in the dependent variable. Since survey respondent are asked to recall their food security status for past months, we believe this is a reasonable choice, thus our preferred specification is to use probit. In any case, our estimates do not vary in any large way based on estimation technique.

We estimate the following equation

$$\Pr(Y_{ict} = 1) = \alpha + \sum_{g=-20}^{20} \beta D(t_i - T_c^* = g) + z_{ict}\gamma + x_t\phi + w_{ct}\tau + \varepsilon_{ict}$$

i indexes individual, c indexes county, t indexes time by month and year.

D is an indicator equal to one in the year and month that a county introduces the EBT as the delivery system for food aid with β being the related estimate of the treatment effect. Naturally this is our variable of interest and reports the impact of EBT card reforms on food insecurity in any month within the 20 month time window on either side of the conversion month. It should be noted that β is relative to the excluded time periods that occur outside the 20 month time window. t_i is the year and month when the interview was

conducted for each individual, and T_c^* is the year and month of conversion for county c. z_{ict} represents a column of individual level characteristics obtained via the CHIS sample. Characteristics included are: age, sex, marital status, household size, presence of children in the household, employment status, housing status, race, a WIC recipiency indicator and health insurance status. α is our intercept term. w_{ct} provides county level characteristics such as mean income, employments rates. mean benefits levels, an urbanicity measure, obesity rates, and whether benefits delivery was mail based or in-person. x_t controls for time varying state food stamp program characteristics. This includes controls for asset limits, and qualifying income thresholds. ε_{ct} are unobserved county, time and individual effects. Finally, Y_{it} is our dependent variable, a dummy for an affirmative answer to having experienced food insecurity.

Again, the variable of interest is β the estimate of the treatment effect. Essentially this tells us the impact of EBT card reforms on food insecurity for a 40 month time window surrounding the month of conversion in each county. Specifically, each observation is given an event time relative to the date of conversion in that particular county. We are then able to consider the average impact of EBT card reforms, over time, for each of the 58 California counties. The framework is attractive in that we are able to observe if there is a treatment effect, but we are also to observe how a treatment effect may vary over time. This differs from the two period, DID framework where the treatment effect is viewed as a one time permanent shift. For a two period model, the treatment effect is effectively the average change for all the periods following implementation of the reform (usually referred to as the post period). The event study framework allows us to see what the treatment impact is

for each individual month, and not just the average of all the months following the reform. Moreover, the event study framework is also attractive in that it allows us to see if there is any change prior to implementation of the reform. Finally, a plot of the treatment effects for each month allows us to have an additional visual gauge for how much underlying noise is present in the estimate.

A result consistent with EBT card reforms altering food insecurity, would be for the treatment effects to statistically indistinguishable from zero for the period leading up to the month in which EBT card reforms take effect, and then experiencing an abrupt change at the date of conversion, followed by months of a sustained shift. This would be consistent with reforms altering food insecurity and having a lasting impact. Another possibility is that EBT cards have an impact, but it is not immediate. This would be shown by the same pattern as just described, except that there may be a lag between intervention and a shift in the magnitude of the treatment effect. Thus, an impact, but not an immediate one. Another possibility, is that EBT card reforms have an impact, but the impact is not permanent. This would result in a shift of treatment effects' magnitude at the moment of intervention, but that the treatment effect will then decrease and return to it's pre-reform mean.

2.5.3 Results

The fundamental question to this study is "Did you eat less than you should because of money?", which we believe is the overall effect of EBT conversion on food insecurity. The simple reason is that this is the most flexible and comprehensive, in the sense that other questions are less general, and therefore they may not capture as complete

an image of the nature of food insecurity. Thus, the estimates from other questions are likely not as representative or general as "Did you eat less than you should because of money?". Since the other questions can be thought of as being ancillary to this question, we interpret EBT's impact on answers as the overall treatment effect on food insecurity. Table 2.2 shows regression estimates for this particular question using multiple specifications. We vary inclusion of controls for year effects, month effects and individual characteristics. The treatment effect estimates are denoted in the following way, t_6 represents the impact on food insecurity of being at a point in time 6 months prior to EBT conversion, t_2 represents the impact on food insecurity of being at a point in time 2 months prior to EBT conversion, t0 represents the impact of EBT card reforms in the first month of EBT card reforms, and t6 represents the impact of EBT card reforms 6 months after the initial reforms.

Table 2.3 presents estimates for the above specification with 4 different dependent variables. The 4 dependent variables are positive responses for the following questions: "Were you unable to afford balanced meals?", "Did you run out of food because of money?", "Did you skip meals in order to conserve food?", and "Did you eat less than you should because of money?" Respectively these are food insecurity measures 1, 2, 3, and 4, as reported in Table 2.3. As stated previously, we believe "Did you eat less than you should because of money?" is the overall treatment effect of EBT conversion of food insecurity. However, we also believe that the other questions are well within our topic of interest, and would like to utilize all the relevant questions to analyze the impact of EBT conversion on food insecurity through multiple angles. Table 2.3 shows estimates for our other three questions of interest. obtained using probit estimation methods following Horrace and

77

Oaxaca (2006).

In general, we find that EBT cards did not have a meaningful impact on food insecurity in the hypothesized way. Recall the hypothesis was that EBT card reforms would diminish food insecurity by reducing the harmful effects of loss and theft, and that by making food stamp fraud riskier fewer people would choose to traffic food stamps. Those people that used to traffic and now do not, would likely have a reduced risk of food insecurity. However, treatment effect estimates are not consistent with the hypothesis. Rather, it appears EBT cards resulted in a very short term, but statistically significant increase in food insecurity. This relationship can be seen by noting the estimates of to and t1 across specifications and dependent variable. The estimates are very distinguishable from zero, moreover there is no evidence of an impact in the months prior to conversion suggesting the impact is not created by some unknown factor that roughly mirrors the timing of EBT card reforms. The result is robust to different measures, inclusion of controls, and estimation techniques. Thus, we feel that the result is a credible one, but perhaps not terribly interesting from a policy perspective. A key take away is that the impact is extremely transitory in nature. The impact is identifiable for only one or two months at the 95% confidence interval and returns to the pre-reform mean level of food insecurity. In terms of the magnitude of the impact, we also present OLS estimates as probit estimates do not lend themselves to a very intuitive interpretation of the coefficient. However, our OLS estimate suggests that residing in a county during the month of EBT card adoption increases the probability of experiencing food insecurity by 2%.

It is worth noting that our estimates are for a 40 month time window, 20 months

78

Table 2.2: Impact of EBT Conversion on "Ate Less than Should Because of Money"

Variable Name	Probit FI Measure 4	OLS FI Measure 4	Probit FI Measure 4
t0	0.2019**	0.0298**	0.2759***
	(0.1017)	(0.0151)	(0.1025)
t1	0.1764*	0.0210	0.2085*
	(0.1075)	(0.0138)	(0.1145)
t2	0.1069	0.0050	0.1195
	(0.0989)	(0.0112)	(0.0974)
t3	0.1791*	0.0206	0.1784*
	(0.0992)	(0.0139)	(0.1043)
t4	0.0325	0.0022	0.0668
	(0.0882)	(0.0095)	(0.0937)
t5	0.0324	0.0002	0.0370
	(0.1073)	(0.0111)	(0.1075)
t6	0.0836	0.0106	0.1070
	(0.1015)	(0.0129)	(0.1032)
t_6	0.1227	0.0118	0.1415*
	(0.0720)	(0.0087)	(0.0769)
t_5	0.1311	0.0151*	0.1426*
	(0.0646)	(0.0089)	(0.0773)
t_4	0.0839	0.0140	0.1068
	(0.0761)	(0.0092)	(0.0790)
t_3	0.1271	0.0083	0.1154
	(0.0734)	(0.0094)	(0.0769)
t_2	0.0851	0.0059	0.0646
	(0.0759)	(0.0088)	(0.0781)
t_1	0.0316	-0.0004	-0.0144
	(0.0655)	(0.0078)	(0.0686)
constant	-1.7489***	0.0811***	-1.3143***
	(0.058)	(0.0082)	(0.0778)
Number of Obs.	191896	192130	191896
Number of counties	58	58	58
R-Squared	-	0.033	-
Year Controls	Yes	Yes	Yes
Month Controls	Yes	Yes	Yes
Individual Characteristics	No	Yes	Yes
Mean Dependent Var.	0.0571	0.0571	0.0571

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

FI refers to Food Insecurity. Thus each of the dependent variables has its origins as a direct question regarding food insecurity as part of the CHIS survey.

Table 2.3: Impact of EBT Conversion on Food Insecurity

Variable Name	Probit FI Measure 4	Probit FI Measure 1	Probit FI Measure 2	Probit FI Measure 3
t0	0.2759***	0.2026**	0.2000**	0.2757**
	(0.1025)	(0.1038)	(0.1026)	(0.1144)
t1	0.2085*	0.0655	0.1704*	0.1958*
	(0.1145)	(0.1075)	(0.1011)	(0.1108)
t2	0.1195	-0.0190	0.1008	0.0471
	(0.0974)	(0.0816)	(0.0910)	(0.1043)
t3	0.1784*	0.1187	0.1965**	0.1829
	(0.1043)	(0.0905)	(0.0911)	(0.1118)
t4	0.0668	0.0461	0.0710	0.0456
	(0.0937)	(0.0910)	(0.0783)	(0.0942)
t5	0.0370	0.0474	0.1063	0.0018
	(0.1075)	(0.0926)	(0.0951)	(0.1140)
t6	0.1070	0.1122	0.1527	0.1078
	(0.1032)	(0.0876)	(0.0823)	(0.1153)
t_6	0.1415*	0.1045	0.1439*	0.1085
	(0.0769)	(0.0688)	(0.0778)	(0.0758)
t_5	0.1426*	0.1647**	0.1255*	0.1362*
	(0.0673)	(0.0576)	(0.0716)	(0.0741)
t_4	0.1068	0.0134	0.0394	0.1145
	(0.0790)	(0.0591)	(0.0554)	(0.0818)
t_3	0.1154	-0.0043	0.0316	0.0520
	(0.0769)	(0.0670)	(0.0652)	(0.0795)
t_2	0.0646	-0.0049	0.0120	0.0609
	(0.0781)	(0.0577)	(0.0661)	(0.0794)
t_1	-0.0144	0.0211	0.0563	-0.0013
	(0.0686)	(0.0661)	(0.0635)	(0.0720)
Constant	-1.3143***	-1.0125***	-1.0392***	-1.4101***
	(0.0778)	(0.0632)	(0.0582)	(0.0812)
Number of Obs.	191896	191837	191813	192069
Number of counties	58	58	58	58
Year Controls	Yes	Yes	Yes	Yes
Month Controls	Yes	Yes	Yes	Yes
Individual Characteristics	Yes	Yes	Yes	Yes
Mean Dependent Var.	0.0571	0.0995	0.0968	0.0518

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

FI refers to Food Insecurity. Thus each of the dependent variables has its origins as a direct question regarding food insecurity as part of the CHIS survey.

prior to conversion and 20 months after. However, for reasons of space, both Table 2.2 and Table 2.3 show estimates for only 12 total months. Importantly, the estimates are obtained from running a regression over a 40 month window, not a 12 month time window. Figures 2.4 through 2.7, graphically reveal the treatment effect for the full 40 month window. We experiment with multiple time windows, however the result does not change in any dramatic way. Moreover, given the timing of the reforms and CHIS data collection, 20 months is the largest practical pre-intervention period. Thus, for the pre- post- period to be balanced we use a 40 month window. We are, however, able to trace out treatment effects for a longer post window, but nothing of interest or value appears to be revealed.

Figure 2.4: Eat Less than Should Because of Money

Figures 2.4 through Figure 2.7 show a graphical representation of the treatment effect over a 40 month time window, 20 months prior to EBT card introduction and 20 months after the initial implementation of EBT card reforms. Figures are obtained using the full set of controls and a probit estimation. The figures offer some value above and beyond

81

Figure 2.5: Food Didn't Last Because of Money
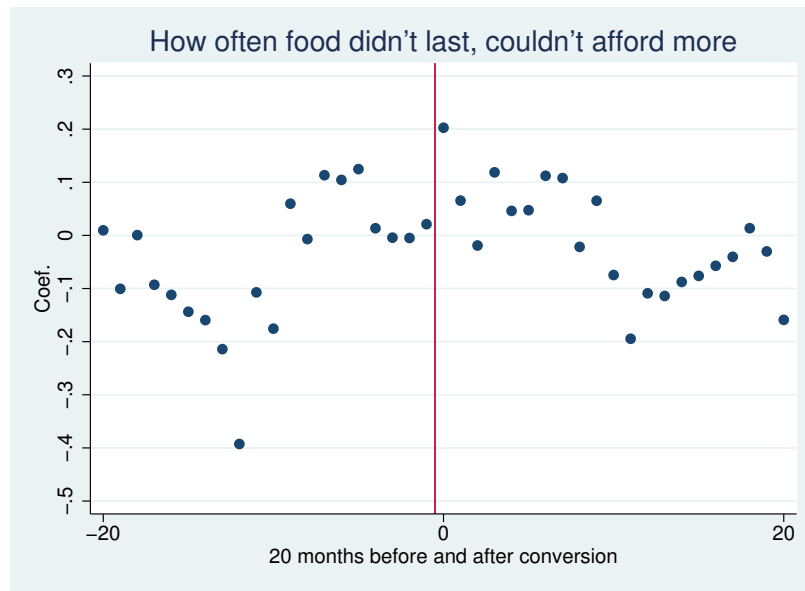
the tables in that we are able to see a larger time window, and obtain a more intuitive sense of how the treatment may have impacted food security over time. Again our results suggest a short lived, transitory increase in food insecurity that occurred following EBT card reforms' introduction. The magnitude varies slightly, but is roughly 2%. Although this is not a massive change, it is notable and very distinguishable from zero, and moreover, the result shows considerable degree of robustness through changes in the measurement of food insecurity. The effect mostly only lasts for 1 month, but in the case of our primary question of interest, "Did you eat less than you should because of money?", the effect is slightly longer persisting for 2 months. This is somewhat expected, relative to other questions, given the more inclusive nature of the question. It's magnitude is also slightly larger, for the same reason, i.e. the more comprehensive nature of the question.

Figure 2.6: Couldn't Afford to Eat Balanced Meals

How often couldn't afford to eat balanced meals
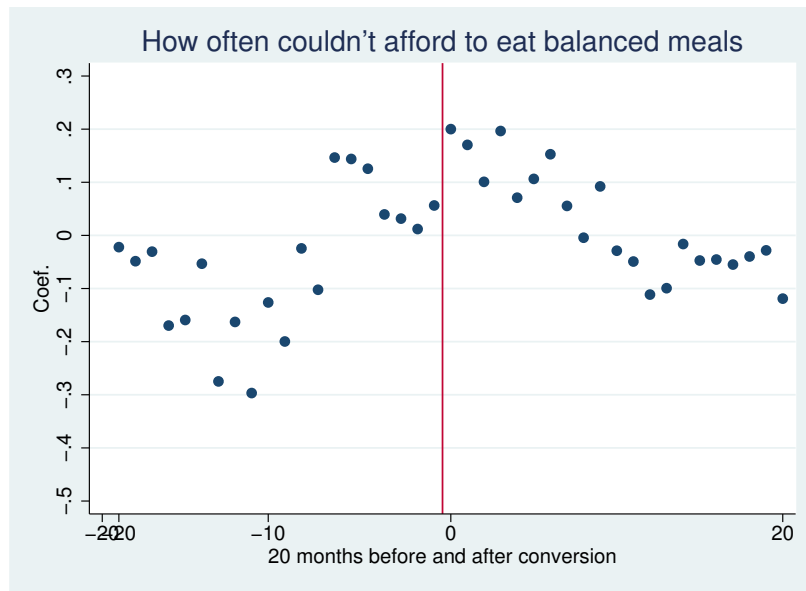
2.5.4 A transitory increase

Interpreting the result requires some thought, in that it is fairly different than hypothesized. Our conclusion is a fairly mundane one, but one we think is quite likely. That is, EBT card reforms did not alter long run food insecurity, but reforms were not a purely smooth process that resulted in a fleeting increase in food insecurity. Presumably, many individuals had trouble using their EBT cards in the initial month or months of the EBT card regime, and as a result experienced an increase in food insecurity. This could have manifested itself through a variety of channels. One important one is that there were differences across counties in terms of how EBT cards were given to recipients. In some counties cards were mailed to beneficiaries, while in others beneficiaries had to come into social service offices and pick up their cards in person. Both methods have advantages and disadvantages, but importantly neither is impervious from beneficiaries not getting their cards in a timely manner. Ohls and Beebout (1993) note that a commonly cited

83

Figure 2.7: Skip Meals in Order to Conserve Food
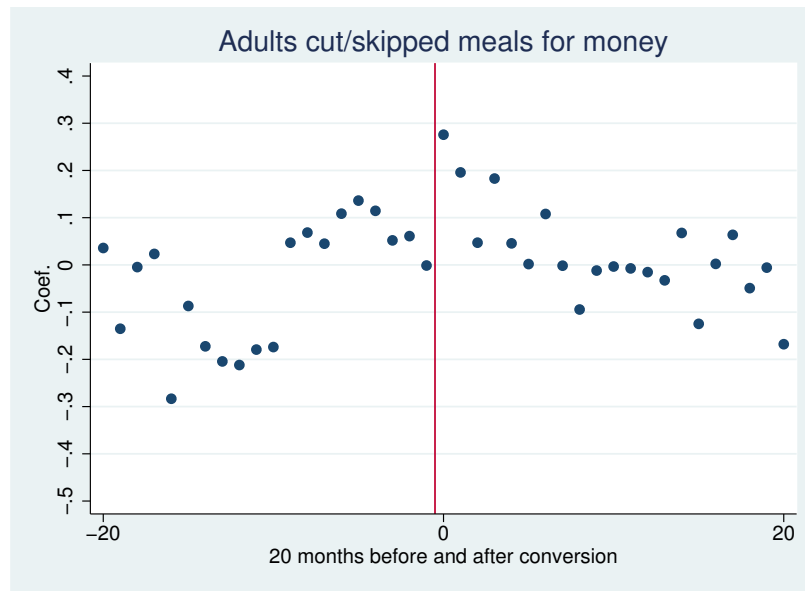
regularity in social work literature is that program participants are often characterized by frequent changes in address that make mail based delivery problematic. On the other hand, mandating that individuals come into the social service agency solves problems associated with mail-based delivery, but is often very difficult for beneficiaries who may not opt to come into the office in a timely manner for a variety of reasons. These reasons would likely include the need to work during business hours, long travel times from home to social service agencies, lack of reliable transport options and the need to coordinate a specific appointment time with their constraints. Finally, at least 6 counties in California used the movement to EBT card regimes to transition from one form of delivery to another. In 5 cases, the county went from mail based delivery to in person pick-up. It is quite likely in these 5 counties many program participants were expecting to receive their EBT cards via mail, but did not. In the one case where the move was from in person to mail based, it is reasonable to expect there to be some difficulties in reaching recipients without stable

addresses.

Another simpler possibility is that some recipients may have had difficulty operating their new cards, or may have found that their cards did not work as intended due to malfunction. In interviews we conducted, county officials have suggested that the transition was fairly smooth and successful, but not completely free of hiccups. For example, Fresno county officials reported that by the second week of EBT card reforms only about half of cards had been used to make a transaction despite the overwhelming majority of the food stamp population already having their cards. Moreover, Fresno county officials stated that currently in a typical month, at least three quarters of food stamp households make a transaction within the first three days of getting a new month's benefits. Although, other counties did not volunteer this particular story, other counties did suggest that the transition was far from painless for many households in the days following reforms.

Another potential problem is that some vendors did not immediately make the transition from paper based to EBT based transactions. This share of vendors is admittedly likely to be quite small, as food stamp customers are a vital segment of the customer base for many vendors. Moreover, they would not be permitted to continue accepting paper based benefits so failure to make a successful transition would result in lost sales opportunities.

2.5.5 Why not a decrease in food insecurity?

All of these possibilities suggest that there are very plausible reasons for why a transitory increase in food insecurity would materialize. A related question then becomes why the hypothesized effect, namely a lasting decrease in food insecurity, did not materialize. Essentially, this reduces to considering the validity of each of the two channels that led to

the hypothesis that food insecurity would fall.

The first channel is that reforms would lead to less fraud, and therefore greater expenditures on food leading to a decrease in food insecurity. The most likely answer is that many of the individuals engaging in food stamp fraud, were able to do so precisely because they were not experiencing significant food insecurity. The thinking being that it would be fairly irrational for a person to traffic food stamps, if doing so would lead to food insecurity.

Although this may appear an innocuous assertion, one of the enduring puzzles food stamp program administers face is that from a purely rational perspective, food stamp trafficking should be a fairly rare occurrence. This stems from the fact trafficking food stamps would only make sense for what are called extra-marginal households, that have ample food supplies. Extra-marginal households are those households that do not spend any of their own income on food. Food stamp trafficking may be attractive for these households as their food needs are met, and they are not selling benefits only to end up having to use their own income to buy food later. Contrast this with what are generally called infra-marginal households. Infra-marginal households are households that spend some of their own income on food. It does not generally make sense for these households to engage in food stamp trafficking, as doing so means that they will have to replace the food they would have gotten with sold benefits. For infra-marginal households, trafficking implies that they sell benefits for cash today, but tomorrow have to buy food and have less purchasing power then if they hadn't sold their stamps yesterday. Since food stamps benefits always sell at a discount, these households would be accepting a loss in real income or purchasing

power. Thus, from a rationality standpoint, infra-marginal households will not choose to fraudulently sell food stamp benefits. However, the majority of food stamp households are considered infra-marginal, and yet food stamp administrators believe fraudulent sales are quite common as noted in Macaluso (2000). Additionally Macaluso reports anecdotal details that suggest fraud is not confined only to extra-marginal households. Thus, the assertion that EBT did not reduce food insecurity via reduced fraud because prior traffickers were extra-marginal and had ample food, is not without doubts.

The second channel is that EBT card reforms would make loss and theft less costly in terms of lost access to food stamp benefits. It is perhaps more challenging to explain why this channel appears to be inconsequential. In theory, it is abundantly clear that loss should be far less problematic in the case of EBT cards. The EBT card itself does not represent benefits, merely a vehicle for accessing an account where benefits are safeguarded. Contrast this with paper stamps, which represent real benefits and it is clear that loss of a card is far less problematic than paper based stamps. Moreover, cards can be replaced within a fairly short period of time. The most likely, but far from definitive, answer would be that loss of benefits did not account for a sizable share of food insecurity. It is still possible that EBT cards did reduce food insecurity through this channel, but since the frequency and/or magnitude of costs associated with loss is quite low and/or the data is too noisy to pick up the effect. The alternative explanation is simply that loss is not an issue for the overwhelming majority of food stamp beneficiaries and regardless of the data we would not pick up an effect.

Theft presents the same considerations, namely that it simply isn't important

and/or that the data contains too much random noise, but with an additional wrinkle. That wrinkle is that theft remains a possibility with EBT cards, but requires that the prospective thief also know the card's PIN number. If it is the case that most thefts are carried out by individuals with knowledge of the PIN number, then we would not expect EBT cards to reduce food insecurity by means of reducing theft. This is simply because the number of thefts would not have dropped since EBT technology has not been effective in making theft more difficult.

A final consideration is the extent to which food insecurity is, or isn't experienced by food stamp program participants. If a large proportion of those suffering from food insecurity are not participants in the food stamp program, this would substantially attenuate our results and increase imprecision in our estimates. Within the CHIS sample fully 36% of those that confirm food insecurity do not report receiving food stamps. This number may be problematic for two reasons. One, is that many who receive food stamps do not admit to receiving benefits as it is potentially shameful. A second consideration is that many who experience food insecurity are not permitted to receive benefits due to matters of immigration status. Moreover, many are not willing to disclose their immigrant status. Thus, there is considerable room for measurement error, as well as an attenuating effect stemming from the potentially large number of individuals who experience food insecurity, but do not claim to receive food stamp benefits.

As to why this is potentially problematic for our analysis, the thinking here is that those experiencing food insecurity, but not receiving food stamp benefits, would not experience any changes in food insecurity as a result of EBT card reforms. Another way

of stating this is that our estimates are not treatment-on-the-treated (TOT) estimates. A potential way around this would be to restrict the sample to only food stamp recipients. However this dramatically reduces our sample size and so thoroughly increases imprecision in our regressions that we are unable to say much of anything of interest. Thus, one of the key shortcomings of the CHIS data has been the relatively small sample size for individuals that report receiving food stamp benefits. Of course, there is likely to be significant under reporting as food stamp program participation is a source of shame for many individuals. Moreover, the CHIS sample reports rates of food stamp program participation substantially less than reported by the USDA during the relevant time window. Thus, there is some evidence for under reporting.

2.6   Conclusion

We investigate the rollout of EBT card reforms in California during the period 1998 to 2006 to determine if EBT card reforms had a causal impact on food security for California recipients of food stamp benefits. Although we expected to see a decline in food insecurity, as EBT cards would likely reduce trafficking, loss and theft of benefits that all contribute to food insecurity, we found no such evidence of a relationship. However, we did find evidence of a very transitory increase in food insecurity immediately following implementation of EBT card reforms. The magnitude of the effect is roughly a 2% increase in the likelihood of an individual reporting food insecurity in the past month. The result is significant at either the 95% or 99% confidence interval and lasts for only one or two months before returning to pre-reform levels. We consider a variety of measures of food insecurity

and find the result fairly robust to changes in measurement and inclusion of controls. We posit that the transitional increase in food insecurity stemmed from program participants' difficulty navigating the new EBT environment. Potential avenues for these difficulties are recipients failure to obtain their EBT cards in time, lack of knowledge about how to use the card, card malfunction, and vendors that were slow to make the transition to EBT transaction terminals.

Although we find the transitory increase in food insecurity credible and highly plausible, we are surprised by the lack of a long-run decline in food security following EBT card reforms. The effect may simply not exist as hypothesized, or our data resources may not allow for sufficiently precise estimates. In terms of offering a possible answer as to why a reduction in food insecurity does not appear to have materialized, we offer a few thoughts.

A distinct possibility is that the reduction in fraud is not meaningful from a food security perspective. This would likely stem from the possibility that the majority of those that traffick in food stamps are extra-marginal, and have distorted consumption as a result of the in-kind transfer. These households do not need to spend any income on food and any excess food benefits are likely to be sold. Whitmore (2002) argues exactly this and Cunha (2014) finds strong evidence for this in international settings. If it is indeed true that only extra-marginal households engage in food stamp trafficking, it is highly unlikely that EBT card reforms would have diminished food insecurity through this channel alone. This is because EBT card reforms would only effect households that were not food insecure to begin with as food insecure households would necessarily be infra-marginal. Although this seems quite likely in retrospect, it is important to consider that an alternative framework

emphasizes differing marginal utilities of cash and food at different times in the month. Under this framework, it is still possible infra-marginal households will choose to sell food stamp benefits. If this were true, we suspect a reduction in food insecurity would be apparent. Thus, our results may be interpreted as supporting the conclusion that infra-marginal household trafficking is a rare occurrence within California.

In terms of the loss and theft channel, we are generally puzzled by the result that EBT card reforms have not been meaningful. Our weak inclination is that, perhaps food insecurity stemming from loss and theft were considerably less than we had previously thought. At present, we are unable to find data to confirm or reject this, but hope in the future such data will be available. We believe this may be a fruitful venue for further research. In particular, the USDA may have data that would shed light on the matter.

Chapter 3

An Empirical Examination of

Divorce in China

3.1   Introduction

A distressing aspect of many nations' transition to modern, affluent societies has been the perceived deterioration of successful marriages. Although the definition of a successful marriage is far from clear, it does appear to be a widely held belief that divorce is not a socially desirable outcome. In many newly industrialized nations, particularly in East and Southeast Asia, increases in economic development, education, and career opportunities have come with an increase in the proportion of marriages ending in divorce. This has been true in the cases of Taiwan, Hong Kong, Singapore, South Korea and Japan. In more recent years, mainland China has also come to exemplify this trend; that is, as China has emerged as a modern and increasingly affluent nation, a common occurrence is that a

great many marriages end in divorce. A litany of explanations have come forward in order to better understand this phenomenon, but for the most part these explanations are based largely on conjecture and anecdote. Although a causal explanation is beyond the scope of this paper, we hope to explain which factors are most associated with divorce, and how they have changed over time for different generations of Chinese citizens.

The increasing occurrence of divorce in China, may be of considerable interest for a variety of reasons. Perhaps most obvious, is that as the world's most populous nation, and one where marriage is very much the norm, there are tremendous numbers of individuals that experience the costly and potentially pernicious effects of failed marriage. These effects can take many forms, among many others emotional, economic and health outcomes are all altered in the face of failed marriages. Another cause for interest, is that as a developing nation the Chinese experience may offer the opportunity to understand divorce in a way that is particularly relevant to other developing nations, particularly those that are on the cusp of industrializing. That is, there may be considerable external validity or generality in reaching a greater understanding of divorce in China.

Moreover, there is considerable evidence that marriage may be desirable for a variety of economic and social reasons. General findings from previous research suggest that marriage represents a valuable mechanism for risk sharing and may substantially increase living standards for marriage partners through the sharing of household expenses and non-rivalrous goods. Moreover, there are significant gains to be had from within household task specialization and consumption complementarities. Perhaps most importantly, there is considerable evidence suggesting there are benefits to child rearing within marriage. This

may be particularly pertinent in the Chinese case as 74% of marriages result in children. Wang et al (2007) find that some of the benefits of within marriage fertility and child rearing (as opposed to child rearing outside of marriage) in China include better infant health, better child educational attainment, better nutrition and better school attendance. Thus, marriage remains a very important family institution that potentially sets the stage for better lives for those children born within the confines of marriage. Additionally for adults, marriage is associated with greater rates of home ownership and greater labor force participation, which presumable offer benefits to society at large.

It is worth noting that it is quite possible that divorce in China may be very different from divorce in modern Western nations in a variety of ways that make this study of greater interest. Since China is a relatively poorer nation, the economic dislocations associated with divorce may be potentially quite a bit more dramatic and costly for those that divorce and must return to life on a single income. Along the same vein of thinking, Wang et al (2009) notes that laws meant to protect those that are economically ill-prepared and not at fault for contributing to a divorce, are relatively less stringent and poorly enforced than in many Western nations. Moreover, since such a high percentage of Chinese marriages result in marriage, it is far more likely a divorce will occur in the presence of children than in Western nations. As such, the impacts of divorce will on average affect children at a greater rate than in the West. Another key difference, is that as a rapidly developing nation the key contributors and outcomes associated with divorce are far less likely to be as stable as in modern western nations. Rather, an accurate understanding of marriage and divorce is more likely to be in a state of flux and change over time that may merit relatively more

frequent investigations that for the relatively more mature economies of Western Europe and North America.

We conduct two related, but different analyses of divorce in China. Our primary objective is to determine which factors are associated with divorce in the Chinese context over time. This would presumably be of interest because the Chinese divorce rate has grown quite rapidly in the last 10 to 15 years. Secondly we wish to determine what impacts might be associated with divorce. In particular, we consider matters of physical health and mental health since an empirical regularity in Western studies has been that divorce is costly from a health perspective. However, this has been largely unconfirmed in the Chinese case. We use two nationally representative data samples, the China Health and Retirement Longitudinal Survey, (CHARLS) and the China Family Panel Study (CFPS). These data sets allows us to obtain a fairly complete understanding of the predictors of divorce, and how the ever divorced population differs from the married, but never divorced population. That is, the dataset allows us to consider how those that experience divorce, are quantifiably different from those that have marriages that do not end in divorce. This would presumably be of interest because the Chinese divorce rate has grown quite rapidly in the last 10 to 15 years and has led to a preponderance of potentially faulty explanations. A clear benefit of having two large, detailed surveys is that it allows us to cross check results from one survey to another, and form a more credible perspective on the predictors of divorce.

Having a more empirically founded understanding of divorce may be quite useful from a policy perspective in that the Chinese government has expressed considerably concern about the perceived break-down of family structures. Although divorce and marriage are

not the only visible signs of family deterioration, they are highly visible and of considerable concern for a society greatly concerned with stability and harmony. If policy makers are better aware of which predictors are strongly associated with divorce policy can be directed in such a way to discourage what might be perceived as inherently risky marriages and encourage those marriages seen to be less likely to end in divorce. For example, if it were found that education was associated with lower rates of divorce it is possible current valuations of the benefits to education are too low. In this case, policy makers would potentially choose to increase their valuation of the positive externalities associated with education

We find that factors more strongly associated with divorce are, several proxies for possessing modern attitudes about divorce such as use of the lunar calendar and viewing Mandarin competence as highly important, having migrated, and feeling one is relatively wealthier than others. The role of migration is not entirely clear. Migration is a strong predictor of divorce for individuals in the CFPS, but is virtually meaningless for the older individuals that are surveyed in the CHARLS. Interestingly, the following factors were generally not associated with a greater incidence of divorce: educational attainment, lack of siblings, income and age at time of marriage. We also consider how both physical and mental health outcomes differ for the divorced, the remarried, and those that marry but never divorce. In general, we find divorce is harmful from a mental health perspective. Some portion of this damage is eliminated via remarriage, but not all as the remarried are generally worse off than the married, but never divorced segment of the population. In terms of physical health, we find some evidence that divorce is associated with worse

outcomes for chronic disease and the average number of hours an individual sleeps. It is important to be clear that our results are not causal, and no attempt is made to present them as such.

The rest of the paper is organized in the following manner: section 2 considers relevant contributions from the Economics literature, section 3 lays out the data sources, section 4 offers background information about the environment of marriage and divorce in China, section 5 offers analysis of the data and results, section 6 offers a short discussion and conclusion.

3.2 Literature Review

The vast majority of research conducted looks at divorce in modern Western nations and considers a variety of explanations, both social and economic, for increases in the rate of divorce. Most studies have been done in the US, although many others also been done in Europe. Relatively fewer studies have been done in the developing world. Moreover, very few English language studies have been conducted with regards to the newly industrialized nations of East and Southeast Asia.

A great many papers from the Economics literature looks at marriage and divorce as being products of rational thinking where individuals choose to marry if the perceived benefits of marriage outweigh the costs of marriage. These papers are discussed in greater detail below, but the central message across nearly all papers is that divorce ensues when the costs appear greater than the benefits. In general, benefits arise from intra-household specialization (particularly in the form of one partner specializing in work within the home,

and child rearing), risk sharing, and consumption complementarities. Costs stem from lost employment opportunities, and loss of individual autonomy. Gary Becker's 1981 seminal paper on marriage laid out the theory of production complementarities as being a key driver of marriage and divorce.

However, US entry into marriage has declined since the 1950s and 1960's. Furthermore, divorce rates rapidly accelerated from the 1960's into the 1980's, but have since stabilized. General explanations for this phenomenon within Economics circles is that intra-household specialization has decreased and become less important. Moreover, some of the benefits of marriage may have decreased in recent years as child rearing within marriage has declined, more production within the home can be replaced with inexpensive purchased substitutes, the legal cost of divorce may have declined, and the need for risk sharing may have also declined. Furthermore, as the labor market has become more open to women and more remunerative, and within the home labor has become relatively easier due to technological advance, the relative cost of staying married or entering marriage is likely to have shifted for many women.

As a general statement, the Economics literature suggests that the relative decline of intra-household specialization is at least partially overcome by the relative increase in the importance of consumption complementarities. The central theme behind consumption complementarities is that marital partners are able to achieve higher levels of utility and leisure because marriage partners are able to share non-rivalrous goods and reduce costs by cohabiting. As partners are better matched and enjoy each others company to a greater degree, consumption complementarities expand.

Below we consider several of the subsets of the Economics literature and how they relate to marriage and divorce.

3.2.1 Household Technology

One of the key reasons for entering marriage and staying in a marriage is the perception that there are gains from specialization within the household. Although it is not necessarily the case, this is most often seen as one partner, usually female, specializing in home production. The other partner specializes in market production, or more explicitly works for a wage. However, in the past 100 years or so household technology has substantially changed, and lowered the level of labor input for most domestic work. Greenwood, Seshadri and Yorokoglu (2005) document this change and suggest housewives were freed from lives of toil and converted into "managers" of an "army of household appliances". An important consideration is that the time savings afforded from these appliances may simply lead to greater consumption of home produced goods (a substitution effect), or may lead to greater leisure and/or consumption of non-home produced goods (an income effect). The authors find that time savings generally translated into greater consumption of non-home produced goods (an income effect), and substantially greater female labor market participation. Although they do not empirically investigate the impact on marriage and divorce, this would alter the relative cost and benefit ratios for those considering entering into or exiting a marriage. More specifically, this trend would diminish the gains from intra-household specialization, as wages could increasingly substitute for home production.

A similar consideration is that technological changes may have contributed to goods that were previously produced at home, now being produced to a much greater

extent in the market. Cutler, Glaeser, and Shapiro (2003) find that during the period 1950 to 1990, time spent on home food production and clean-up was roughly halved. Moreover, the primary explanation for this decline was the greater availability of market produced food, and the decline of home produced food. Again, this would seem to diminish the gains from intra-household specialization.

This is surely an avenue of thought that would be highly pertinent to China in the modern era, but perhaps not so much in the most recent years when divorce rates have grown the most. Prior to modernization China was very much an economy with considerable home production and has since moved to a new paradigm with dramatically less home production. However, it is quite likely that these changes have diminished in recent years as China as converged to modern market based economy. Thus, this may be an important consideration, but not likely a sole determinant.

3.2.2   Skill-Biased Technical Change

Stevenson and Wolfers (2007) point out that the change in the relative benefits of single life versus married life, is not likely to be felt uniformly across the population. Specifically, technological changes will be far more beneficial to those that have skills well suited to the labor market. Moreover, it is relatively less skill intensive to use a washing machine than to use older laundry technology. To the extent that this example is representative, the relative value of home production may decline. As more individuals relocate from home production to market activity, selection effects will favor those that are relatively higher skill. Therefore, it is not surprising that women with more market oriented skills had relatively less to gain from marriage, and were therefore less likely to marry (or more likely to

divorce). The authors note that only now that most would characterize women's entry into the labor market as being *en masse*, do we find that the education gap in marriage rates among women has closed. Presumably this is driven by the fact that until quite recently, well educated women had less to gain from marriage and were therefore less likely to enter into marriage. However, reverse causality remains a possibility as those women less well suited to marriage may pursue greater educational attainment, which results in fewer gains from marriage.

Since there is little evidence for decreased entry into marriage within China, this would likely manifest itself in greater divorce rates for those that are better suited to the labor market. Educational attainment would likely present itself as a credible proxy for those with greater skills and those that are better suited to labor market participation.

3.2.3 Consumption Complementarities

In essence, consumption complementarities suggest that marriage partners may be able to enjoy leisure and consumption to a greater degree with one another, than without each other. This primarily stems from similar tastes and preferences that arise around non-rivalrous goods, as well as positive externalities that arise from shared consumption. Moreover, many goods and services may be able to be consumed at greater levels in a shared framework, than in an independent one. Lam (1998) suggests that marriage may be increasingly built around this aspect of marriage as the gains from intra-household specialization decline.

An interesting feature of Becker's 1981 work is that in a marriage built around production specialization, there tends to be assortative matching in the sense that one

partner will typically be a high wage earner, while the other partner will be a low wage earner. However, Lam suggests that as consumption complementarities become relatively more important, this type of assortative matching is likely to decline as consumption complementarities are likely to be greater for those of similar, educational, skill and income backgrounds. It is also noted that in theory, assortative matching should persist in a risk sharing form for two career marriages. The thinking here is that marrying someone unlikely to experience unemployment at the same time as you, may be a valuable method of risk sharing or consumption smoothing.

It is not entirely clear how the relative importance of consumption complementarities may affect the likelihood of a marriage ending in divorce. It is quite likely that the relative decline in production specialization and the rise in consumption complementarities will alter the individuals that will select into marriage. Since the population entering into marriage is different, it is likely that explanations of divorce and it's relative frequency will change. However, it is not clear if in theory we should expect more, or less divorce. Although the theoretical impact of increases in consumption complementarities is ambiguous, we thing that the issue may be relatively pertinent in the Chinese case since there is very little that is potentially specific to Western nations about the theory. Moreover, consumption complementarities may be particularly salient in an environment of rising consumption standards. Surely China during the past 25-30 years is a case of dramatically increasing consumption standards.

3.2.4 Changes in the Legal Environment of Divorce

A fairly active area of research in the US has been the effects of changing divorce laws on divorce. Namely, laws have changed to allow for unilateral divorce and no-fault divorce, but there has been significant variation in the time implementation of these changes, as well as spatial variation, in that most divorce laws in the US are determined at the state level. Becker et al (1977) argue that a move towards no fault divorce should not in theory alter the frequency of divorce, because a change in the laws represents a shift in property rights from one spouse to another. As such, they apply Coase Theorem and argue that the end result should be relatively unaltered. However, this argument has not been widely accepted for the simple reason that Coase theorem is not likely to apply to the case of divorce where costless bargaining is likely the exception, and not the rule.

In the case of unilateral divorce, Wolfers (2006) examines divorce rates and laws across states during the period 1956–1998. Findings suggest that divorce rates rose sharply in the two years following the adoption of unilateral divorce laws, but subsequently the divorce rate reverted back toward earlier levels. A decade after these reforms no discernible effect on the divorce rate remains. This is explained as resulting from an interpretation of the courts as servicing pent up demand. Once it has been serviced, the divorce rate reverts to its old pre-unilateral divorce rate.

An older literature investigating unilateral divorce exists and includes contributions from Peters (1986 and 1992), Allen (1992), and Friedberg (1998). Although different conclusions are reached as to the nature and presence of an effect, there are some points of significant agreement that bear consideration. Namely, that any effect is either small or

non-existent, and the presence of an effect does very little to explain the large increases in US divorce rates post 1960.

A different consideration is that unilateral divorce may change the nature of intra-household bargaining. Stevenson and Wolfers (2006) find that female suicide and domestic violence fell following passage of unilateral divorce laws, suggesting a shift of bargaining power to women in marriage. This may alter the choice to enter into marriage in the first place. Rasul (2006) finds that indeed, the presence of unilateral divorce laws contributes to a decline in the marriage rate.

In the case of China, the legal environment has not changed until 2014. Since these changes have occurred so recently they are not within the scope of this paper. However, while we generally feel the legal environment is an issue of considerable importance in explaining marriage and divorce. We feel that in the case of China, the legal environment has been fairly static while divorce rates have not. Thus, we do not expect legal matters to be a key determinant in explaining the recent surge in divorce.

3.2.5 Changing Labor Market Compensation

Since the 1970's an empirical regularity in the US has been rising wage inequality between high and low skill workers. Several explanations have been considered for explaining this, such as skill biased technical change, de-unionization, and declines in the real value of minimum wage work. However, from a theory standpoint, increased wage inequality may have interesting impacts on marriage and divorce. Gould and Paserman (2003) argue that increased wage inequality has increased the value of a good marital match, thus prolonging the search process and resulting in fewer matches. Essentially, greater wage inequality in-

creases the option value of remaining unmarried and causes greater valuation for continuing in the search process. This is tested at the city level by comparing cities over time with different levels of wage inequality. The central finding is that in cities typified by greater inequality, the authors find marriage rates decreased by as much as a third for young women.

Blau and Kahn (2000) also consider wage inequality, but from a gender perspective. They argue that explicitly sexist hiring, management and compensation policies have diminished over time, and the resulting decrease in wage inequality increases the opportunity cost of marriage for women as a result of diminished gains from intra-household specialization. Blau and Kahn do not, however, consider the matter from an empirical perspective.

A common theme in marriage across countries and cultures is that typically men enter marriage at a later age than do females. Becker, Landes and Michael (1977) suggest that this is because the specialized investments women made in preparation of being a homemaker, were less valuable to singles. Moreover, these specialized investments in human capital were typically made at a younger age then the investments men made in preparation for entry into the paid labor market. Thus, men were more likely to enter into marriage relatively later than females. However, in recent years this trend has declined and in many developed nations there is near parity in age at first marriage.

Caucutt, Guner, and Knowles (2002) posit that as labor market returns to the highly educated increase, both sexes, but especially women, will find it advantageous to delay marriage to later age. Although it is not clear what impact this may have on divorce, the authors suggest it may result in better matches, but may also decrease within mar-

riage child bearing, which may alter divorce decisions. Presumably better matches should decrease divorce, but decreased marital fertility may increase divorce.

A general problem with this vein of work is that there exists a considerable possibility of reverse causality. Stated simply, women that are less enthusiastic about marriage or believe they are not well suited to successful marriage are very likely to make greater investments in their human capital, and actively seek better labor market outcomes. Similarly, they are less likely to make investments geared towards maximizing the gains from within household production specialization. Prominent examples of research illustrating these possibilities are Johnson and Skinner (1986) and Stevenson (2006), which respectively show that women who anticipate divorce are far more active in the labor market, and unilateral divorce laws led to greater labor market participation for both married and unmarried women.

This strand of thinking may be quite valuable in the Chinese case, in that women's labor market participation has continued to increase in recent years. Women's compensation has also increased as part of a general trend of rising wages, as well as rising educational attainment among women. However, increases in women's labor market participation have been modest in recent years, while divorce increases have been decidedly immodest. As such, it is likely rising wages and labor force participation are important, but not sole contributors to increases in divorce.

3.2.6 Emergence of reliable, low cost, birth control

Many academic papers have considered how the birth control revolution ushered in by widely available oral birth control may have affected marriage, divorce and fertility.

106

Akerlof, Yellen, and Katz (1996) argue that the availability of reliable birth control and abortion, put women under greater pressure to engage in pre-marital sex, but left them less able to extract a promise of marriage in the event of pregnancy. They document a notable increase in out-of-wedlock births from the 1960's through the 1980's and argue that this reflects a movement away from marriage, and subsequently may have resulted in fewer divorces as well. Bailey (2010) uses a Supreme Court decision that struck down laws banning sales of contraceptives to empirically identify the impact of oral birth control on fertility. In this case, findings suggest that previous work substantially underestimated the impact of oral birth control on fertility, and contributed to a reduction in marriage. Katz and Goldin (2002) make the case that oral birth control likely delayed marriage and reduced fertility by reducing the number of fertile years a woman is married. A key point worth noting is that the decline in fertility would decrease the degree to which intra-household specialization justifies marriage, in the sense that a significant benefit to this specialization is child rearing. In the absence of children, or in an environment of decreased fertility, the gains from marriage would decline. This could manifest itself in either lower marriage rates, higher divorce rates, or both (which is what is observed in the time period).

This may be of interest in the Chinese case because Chinese marriages are different than Western marriages from a fertility perspective in two key ways. One is that a relatively high percentage of Chinese marriages result in children with estimates varying over time from 84% to 93%. This is noticeably higher than in Western nations. However, since the 1980's the overwhelming majority of married couples that do have children, only have one. Thus, it is not clear whether this would constitute decreased fertility within marriage and

potentially lead to more divorce. However, it is worth noting that these fertility trends have been fairly stable in China during the past 10-15 years when the divorce rate has rapidly accelerated. Thus, is is unlikely that changes in within marriage fertility are a key driving factor in explaining divorce.

3.2.7   International Comparisons

Unfortunately, from our perspective, little has been done to consider how Chinese marriages and divorce are similar to, or different from the experiences of other nations. However, there has been a fair amount written about how the US compares to other industrialized nations and how that has varied over time. Wolfers and Stevenson (2007) find that there are some distinct commonalities between the US and some nations, but that there are very large differences with others. Italy is cited as perhaps being the most similar in that attitudes about the relevance of marriage (very relevant) and cohabitation rates (low) are quite similar. Divorce rates are lower in Italy, but convergence appears to be likely. However, a key difference is that remarriage is much more common in the US, while in Italy it remains somewhat uncommon.

A stark contrast with the US, is what is sometimes referred to as the Nordic Model of Marriage, which is characterized by high rates of cohabitation, low marriage rates, and high levels of extra-marital fertility. Wolfers and Stevenson posit that nations such as Sweden are increasingly adopting the attitude that marriage and fertility need not be linked. They also offer evidence that many other high-income nations such as France, the UK and Canada are moving in this direction, with increasing rates of cohabitation, increasing fertility outside of marriage, later age at first marriage and declining rates of

marriage.

3.3 Data

Data for this investigation stems from two primary sources: the China Health and Retirement Longitudinal Survey (CHARLS), and the China Family Panel Studies (CFPN).

3.3.1 China Health and Retirement Longitudinal Survey

The CHARLS dataset is a nationally representative longitudinal survey of adults age 45 and older. The baseline national wave of CHARLS was fielded in 2011 and includes about 10,000 households and 17,500 individuals in 150 counties/districts throughout the nation. The individuals are followed up with every two years. All data is to be made public one year after the end of data collection, although in reality this has not happened. CHARLS makes use of a multi-stage stratified PPS sampling method. However, it is important to note that at this time CHARLS is not longitudinal in the sense that data from the second wave of collection has not yet been released. Response rates are 86%. The interviews are done by household, where one household member responds to a variety of questions about themselves, their household, other household members, the past experiences of themselves and their household members. Importantly individuals are asked about their previous marriage histories and questions are asked about current and past spousal partners.

The survey is somewhat limited in that its intended respondents are those aged 45 years or more. As such, the sample may not be highly applicable in explaining the recent rise in divorce rates. This stems from the fact that the sample is biased towards an older

generation that divorced at much lower rates than the current generation, which divorces with much greater regularity. On the other hand, it is attractive in that respondents have had ample opportunity to marry, divorce, and in many cases remarry. Furthermore, it is attractive in that it allows us to account for some of the long term consequences associated with divorce, in particular the survey provides a good deal of information regarding the physical and mental health of respondents. We are thus able to consider whether divorce is associated with adverse health outcomes.

3.3.2   China Family Panel Studies

China Family Panel Studies (CFPS) is a nationally representative, annual longitudinal survey of Chinese communities, families, and individuals. The baseline wave was conducted in 2010. The CFPS collects information every two years and focuses on the economic, and non-economic, wellbeing of Chinese children and adults. A fairly wide range of issues are addressed including economic activities, educational outcomes, family dynamics and relationships, migration, and health. The CFPS is funded by the Chinese government and administered and maintained by Peking University. As stated previously, although the survey is a panel it is for our purposes a single cross section as the 2012 survey has not been released to the public yet. The 2010 survey interviewed about 15,000 families, and over 40,000 individuals within these families. Moreover, the survey is meant to be nationally representative as it draws responses from all over China.

The survey is attractive for a variety of reasons. Namely, the survey contains many observations and allows for considerable precision in estimation, the survey asks questions that allow researchers to re-construct an individuals entire marital history, and the survey

110

contains a rich set of background information about each individual.

A shortcoming of both data sets is that they present information only regarding marriage and divorce, as opposed to other outcomes such as separation or estrangement that have not resulted in legal divorce. Thus, a severely troubled or de facto failed marriage will be counted as a successful one, so long as it does not reach the point of legal divorce.

3.4 Background

3.4.1 Historical Trends

According to Diamont (2000) divorce has only become common in China since the nation's development rapidly accelerated following initial economic reforms in 1979. Prior to the founding of the People's Republic of China in 1949, divorce was an extremely rare occurrence and one commonly rooted in infertility. With the communist revolution came some change to the acceptance of divorce, and it became fairly straight forward from a *de jure* perspective to obtain a divorce. However, it remained a fairly rare occurrence, at least partially because divorces required government permission that was difficult to obtain and frequently required years of petitioning. Beginning in the 1980's, the Chinese government became far more willing to grant divorces and divorces became more common, but still far below that of most industrialized nations. Throughout the 1990's divorce rates increased further, albeit at a fairly modest rate. Starting in the early 2000s the rate of divorce dramatically increased and reached a level similar to other industrialized nations in East and Southeast Asia. Figure 3.1 shows China's gross divorce rate over time. Since 2004 divorce rates appear to have stabilized.

Figure 3.1: Divorce Rate in China, 1985 - 2013
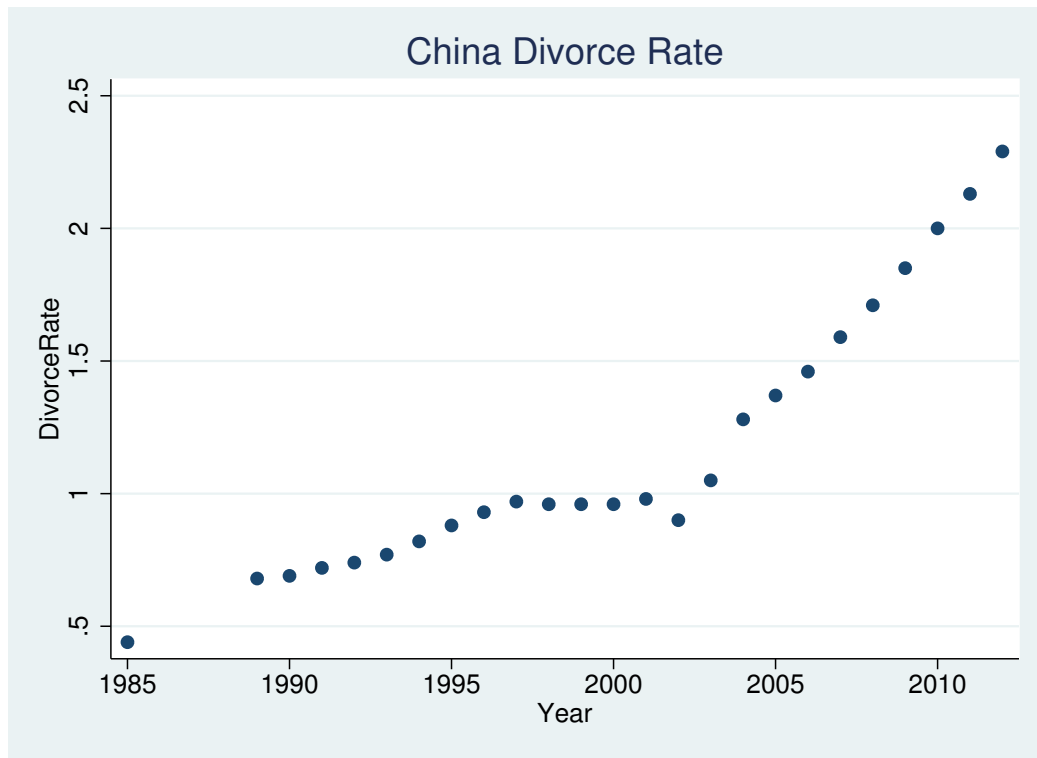
Figure 3.2 shows divorce rates over time for the US, Singapore and China. We chose to show US rates as this is perhaps the most familiar context while Singapore presents highly reliable data using the same metric as the US and China. Moreover, Singapore may be of interest as it is another Asian nation that has experienced rapid growth and has a majority Chinese population from an ethnicity and language perspective. The key take away for our purposes is that while China's divorce rate as increased in a dramatic fashion in recent years, Singapore and the US present different trends. Although the US continues to have a higher divorce rate than either other nation, its divorce rate has actually declined in recent years. Singapore, on the other hand, has been fairly stable. Thus, China might not be viewed as part of a general global increase in divorce rates, rather the increase in divorces in China is likely not explained by global or regional phenomena.

Figure 3.2: Divorce Rates Over Time for US, Singapore and China
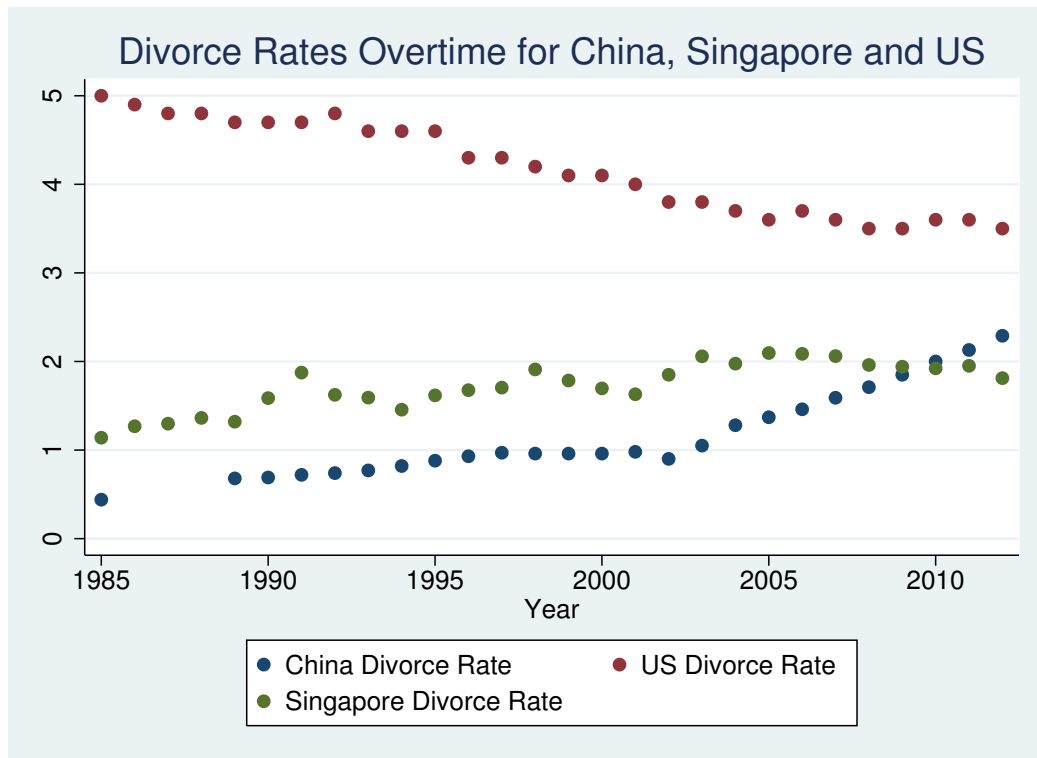
3.4.2 Common Explanations for Divorce Increases in China

As China's divorce rate has dramatically increased, a variety of possible explanations have surfaced. Most are merely touted as explanations within the public sphere, but many of them are very popular and likely believed by a great many people. We hope to shine light on the validity of these explanations.

Increasing economic opportunity costs to traditional marriage

In the case of this explanation, marriage is viewed as being an impediment to economic mobility, namely that by being married, labor mobility is compromised and job seekers are not able to find the same quality of match in the labor market, as they would were they single and unmarried. This argument implicitly assumes that individuals are

113

not fully aware of the opportunity cost imposed on their labor market outcomes prior to marriage. The costs becomes clearer after marriage, and as a result some marriages end. An interesting aspect of this argument is that labor mobility in China has increased dramatically in recent years, and to some extent this has coincided with the increase in marriages ending in divorce. Another interesting aspect of this explanation is that it does not imply that the increase in divorce is necessarily an adverse outcome. Rather, individuals make a rational choice about the trade-offs associated with continuing a marriage in the face of opportunity costs.

One child phenomenon

This explanation has gained considerable traction in recent years and has become a favorite explanation in the Chinese media. The rationale behind this possibility is that the current generation of Chinese young adults marrying and divorcing is composed largely of individuals that were only children without siblings. The thinking is that such individuals are accustomed to being the center of attention, and may be ill prepared to endure the sacrifice and compromise that they would have endured had they grown up in the presence of siblings. In its essence, this argument hinges on the claim that the presence of siblings serves as valuable preparation for marriage, since it may lead to the development of skills to prevent, diminish or resolve inter-personal conflicts. An interesting component of this argument, is that it is perceived to be fairly China specific, as only China has implemented family planning policies that have effectively created an entire generation of only children.

Increased Education

A feature of China's development is that, over time, educational attainment has increased. As a result, the generation of young people marrying today tends to be substantially better educated than the generations that have preceded it. This, of course, roughly coincides with increases in the divorce rate. The intuition is that increased education may make marriage relatively less attractive from a variety of perspectives. This is very much mirrored in the Economics literature that suggests intra-household specialization gains decrease as both partners become wage earners. A point worth noting, however, is that this argument relies to some extent on the premise that education has made continuing a marriage relatively less attractive, but likely has not altered the proportion of individuals entering into marriage, which has remained fairly static over time. This would be necessary to reconcile the increase in divorce with the lack of a decline in the number of marriages entered into.

Adoption of Western Attitudes

This argument posits that traditional Chinese culture is relatively less pre-disposed to divorce than is modern Western culture. As China has modernized and industrialized there has been an increasing influence of Western values and beliefs, regarding marriage and divorce. These cultural changes may have made divorce relatively more acceptable, and as a result more couples opt to end their marriages. From an Economics perspective one might think of divorce as carrying certain shame costs, however, those shame costs are not felt equally across the population. Rather, those with more traditional attitudes, will

feel the shame costs more acutely.

3.4.3 Experiences of other Newly Industrialized Countries

Other industrialized nations in East Asia are also typified by relatively large increases in the gross divorce rate in the past 50 years. Both South Korea and Taiwan have experienced very dramatic increases in their divorce rates and are now among those nations where the divorce rate is the highest. Taiwan's gross divorce rate went from being roughly 0.2 in 1980 to 6.2 as of 2012. This places it behind only Russia, the US, the UK, and South Korea. Taiwan, in particular, may be the best comparison to mainland China as Taiwan possesses roughly the same culture, as both are predominately Han Chinese, and speak Mandarin Chinese as the sole official language.

The experiences of Japan and South Korea are worth considering as well. Although their cultures are far more distinct in relation to China than is Taiwan, there is much cultural overlap and these nations have also experienced very dramatic growth trajectories. In the case of South Korea, the pattern of rising divorce began modestly in the 1980s, accelerated in the 1990s and appears to have peaked in the last 6-8 years. Thus, the experience has been quite similar to that of Taiwan. Japan's economic expansion occurred much earlier than did China, Taiwan or South Korea. It's divorce rate has seen a more steady rise in the past 40 years and does not appear to experienced a period of rapid growth. However, as a result of a near continuous rise in divorce rates, Japan as of 2012 has a fairly high gross divorce rate estimated to be about 2.2 per 1000 people per year. However, it has been suggested that this metric gives an artificially low number since Japan is a fairly elderly society and divorce is primarily an outcome associated with young and middle aged adults.

116

3.4.4 Geographic Context

China is a geographically heterogenous nation from many perspectives, but this is true from a divorce perspective as well. Figure 3.3, shows the divorce rate by province. Notably, there are some pronounced differences by province. Some notable take aways are that the three Northeastern provinces of Liaoning, Jilin and Heilongjiang have significantly more divorce than most of the rest of the country, while the southeastern provinces of Guangdong and Guangxi have significantly less. However, apart from those two areas and the very lightly populated far west, there appears to be fairly similar rates of divorce despite very different levels of economic development.

Figure 3.3: Map of China: 2012 Divorce Rate by Province

Heilongjiang21.96
Jilin21.86
Liaoning22.91
Xinjiang21.64
InnerMongolia11.32 Beijing10.2
Tianjin12.25
Hebei7.18
Ningxia8.14 Shanxi7.12
Shandong6.45
Qinhai16.73
Gansu7.45
Shaanxi7 Henan5.26
Jiangsu8.63
Tibet12.87
Sichuan12.14 Hubei8.91 Anhui7.21 Shanghai13
Chongqin15.24
Zhejiang7.61
Jiangxi6.97
Guizhou10.75 Hunan10.81
Fujian7.2
Yunnan10.04
Guangxi8.0 Guangdong4.65
Hainan5.02

Many have suggested that generally higher levels of development and affluence drive increased divorce rates. However, a simple examination of divorce by region would suggest this interpretation may be problematic. In particular, the relationship is far from perfect as there are many areas where the divorce rate is quite different than what might be expected, given levels of development. If development levels were sufficient to explain

117

changes in the divorce rate, we would expect northern China to be characterized by lower rates of divorce, than is southern China. However, there is not a large or systematic difference between northern China and southern China, and if there is any relationship at all, it would appear to be the opposite of expectations.

3.5 Analysis

We wish to determine those factors the strongly predict an increased likelihood of divorce. Moreover, we are interested in presenting evidence that will either substantiate or weaken many of the potential explanations for rising divorce in China. We also wish to consider how divorce and re-marriage affect health outcomes for the ever divorced population. Finally, we wish to see how the predictors of divorce are different for different generations of Chinese. We conduct separate analyses using the CHARLS and CFPS datasets.

We estimate the following equations using ordinary least squares.

For the CHARLS data set:

$$divorce = \alpha + \beta predictors + \gamma geographic\_controls + \theta economic\_controls + \epsilon$$

For the CFPS data set:

$$divorce = \alpha + \beta predictors + \gamma geographic\_controls + \theta ethnicity\_controls + \epsilon$$

Note that the specifications are somewhat different, but in general fairly similar. Differences in the specifications are motivated by the fact that the two surveys do not contain

the same questions, and therefore present different information sets about the respondents. The similarities are motivated by two considerations. One, is that we want to be able to make as close to direct comparisons as possible so that we can see how potential divorce predictors differ across the two populations in the two surveys. This should allow us to gain some insight into how divorce has changed from the older generation of Chinese, to a younger generation. The second reason is that our choice of independent variables is motivated by the factors we believe to be the most important predictors of divorce. Thus, we choose what we believe is most likely to be the best specification given data constraints.

Geographic controls consist of province dummies, ethnicity controls consist of a series of dummies for different ethnic groups, economic controls account for employment status, retiree status, asset value, and housing square footage. Table 3.1 presents general results. Columns 1 and 2 consider the entire population, but use different OLS and probit estimation techniques. Since we consider a probit specification we include estimates for marginal effects in column 3 directly next to column 2, which reveals coefficient estimates for our probit specification. The results suggest use of one technique over the other is not an important distinction. Columns 4 and 5 consider divorce for the opposing populations of urban and village residents. Columns 6 and 7 consider the differences in the well educated population versus those that are relatively less educated. Finally, columns 8 and 9 consider those with siblings versus those that do not have siblings.

Table 3.1: Predictors of Divorce

VARIABLES	(1) OLS divorce	(2) Probit divorce	(3) Probit ME divorce	(4) Village divorce	(5) Urban divorce	(6) Educ divorce	(7) Non Educ divorce	(8) No Siblings divorce	(9) With Siblings divorce
Age	0.00415	0.158	0.00678	-0.110	0.0183	-0.00143	0.00904	0.0488	0.000596
	(0.0156)	(0.324)	(0.014)	(0.0833)	(0.0147)	(0.0235)	(0.0203)	(0.0954)	(0.0158)
Lunar	-1.09***	-19.6***	-0.941**	-2.26	-0.721*	-1.43***	-0.391	0.997	-1.18***
	(0.378)	(7.16)	(0.382)	(1.59)	(0.368)	(0.502)	(0.579)	(2.29)	(0.382)
No Migration	0.355	6.88	0.279	-0.00585	0.250	0.278	0.525	-4.54	0.560
	(0.457)	(9.26)	(0.355)	(1.76)	(0.454)	(0.609)	(0.694)	(2.91)	(0.461)
Educ Primary	-0.208	-2.89	-0.124	1.76	-0.207			0.923	-0.238
	(0.389)	(8.39)	(0.363)	(2.77)	(0.358)			(2.30)	(0.393)
Educ Middle	0.377	7.58	0.333	-0.213	0.309	0.282		-0.607	0.393
	(0.415)	(8.60)	(0.388)	(2.10)	(0.393)	(0.451)		(2.67)	(0.419)
Siblings	-0.138*	-2.96*	-0.127*	-0.006.90	-0.144**	-0.109	-0.178*		-0.118
	(0.0774)	(1.60)	(0.068)	(0.376)	(0.0734)	(0.111)	(0.105)		(0.0834)
Age First Marriage	0.171***	2.83***	0.121***	0.157	0.167***	0.175***	0.160***	-0.387*	0.199***
	(0.0372)	(0.638)	(0.027)	(0.176)	(0.0354)	(0.0558)	(0.0481)	(0.213)	(0.0377)
SR Wealth	1.04***	19.9***	0.852***	2.62***	0.837***	1.31***	0.736***	0.847	1.03***
	(0.196)	(3.90)	(0.164)	(1.00)	(0.185)	(0.285)	(0.261)	(1.14)	(0.198)
From Village	3.20***	40.2***	2.41***			3.41***	1.30	1.34	3.36***
	(0.521)	(8.48)	(0.676)			(0.614)	(1.20)	(2.77)	(0.531)
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.97***	-340.2***		-1.33	-5.17***	-5.68**	-4.47**	8.06	-5.58***
	(1.66)	(33.1)		(8.52)	(1.57)	(2.36)	(2.23)	(9.99)	(1.68)
Observations	9,351	9,351	9,351	1,047	8,304	5,379	3,972	386	8,965
R-squared	0.115			0.116	0.065	0.0119	0.058	0.120	0.107
Data Source	CHARLS	CHARLS	CHARLS	CHARLS	CHARLS	CHARLS	CHARLS		

SE in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Coefficients and SE are multiplied by 100 for ease of interpretation

3.5.1 Age

It is worth noting that many of the subsequent factors associated with divorce are likely to positively correlate with age. For example, older people are more likely to harbor traditional values, are less likely to migrate for economic reasons and are more likely to have married at a younger age. Moreover, those that are older have on average spent more years as married individuals, and are therefore more likely to have experienced a divorce. As such, it is very important to control for the age of respondents. Moreover, there is the possibility that age may interact non-linearly with both divorce and many of the correlated explanatory variables. For this reason, we consider a wide range of specifications that allow age to enter into the regression non-linearly. We consider using polynomial terms, as well as using dummies for age categories. In both instances, the results of additional explanatory variables is non-significant via an F-test. As such, we have presented results for the most intuitive, but also basic, specification allowing age to enter the regression very simply as just the age of the respondent at the time the survey was conducted.

Table 3.1 and Table 3.4 present results for estimates of the coefficients associated with age. As expected age is highly significant and positive, although this is fairly uninteresting since the cumulative probability of a marriage ending in divorce increases as a couple remains married longer. Thus, the result is not one we view as an important predictor of divorce, but rather, a necessary inclusion to diminish the extent of omitted variable bias. This result and interpretation is true across data sets and specifications.

Table 3.2: Divorce and Related Physical Health Outcomes

VARIABLES	(1) cd1 Chronic Disease	(2) cd2 Chronic Disease	(3) hs1 Hours Sleep	(4) hs2 Hours Sleep
Remain Divorce	-6.05		-10.6	
	(4.08)		(16.9)	
Remarried		4.97**		16.8*
		(2.34)		(9.67)
Age	0.500***	0.496***	-1.97***	-1.98***
	(0.0498)	(0.0498)	(0.206)	(0.206)
Lunar Calendar	5.31***	5.43***	-21.5***	-21.2***
	(1.21)	(1.20)	(4.97)	(4.96)
From Village	2.98*	2.71	-6.55	-7.10
	(1.66)	(1.66)	(6.86)	(6.83)
No Migration	-3.89***	-3.76***	10.1*	10.5*
	(1.45)	(1.46)	(6.01)	(6.02)
Educ Primary	-1.32	-1.33	18.2***	18.2***
	(1.24)	(1.24)	(5.10)	(5.10)
Educ Middle	-1.22	-1.25	2.46	2.39
	(1.32)	(1.32)	(5.44)	(5.44)
Sibling	0.554**	0.566**	-1.34	-1.31
	(0.246)	(0.246)	(1.02)	(1.02)
Age First Marriage	-0.355***	-0.344***	0.399	0.441
	(0.00118)	(0.00118)	(0.00491)	(0.00491)
SR Wealth	4.19***	4.10***	-23.9***	-24.2***
	(0.624)	(0.624)	(2.58)	(2.58)
Constant	30.8***	30.6***	829.1***	828.5***
	(5.28)	(5.28)	(21.8)	(21.8)
Observations	9,277	9,277	9,270	9,270
R-squared	0.125	0.126	0.130	0.130
Data Source	CHARLS	CHARLS	CHARLS	CHARLS

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Coefficients and SE are multiplied by 100 for ease of interpretation

3.5.2 Only Children

A common explanation for the increase in the prevalence of divorce in China is that the current generation of Chinese young adults entering into marriage is a generation largely comprised of only children that were not reared in the presence of siblings. The intuition behind this is that only children may be less adept at resolving conflict and reaching compromise with their spouses as they do not have a history of doing so with siblings. Although we are unable to make any claims as to the validity of such thinking, we are interested in whether or not the presence of siblings is associated with diminished probabilities of divorce.

We examine this from a variety of angles. We first consider what the CHARLS data set suggests. In table 3.1, specifications (1) through (7), we explicitly include number of siblings as an independent indicator variable. We find as a result that, in general, the presence of siblings decreases the probability of divorce. However, the result is only significant at the 10% confidence interval, as such the relationship is not impervious to doubts. In terms of a magnitude, we find that is generally between 0.001 and 0.002. Stated more intuitively the presence of siblings decreases the probability of divorce by roughly 1‰. This would suggest that the presence of siblings may diminish the likelihood of divorce, but the relationship is quite weak. Thus, at least for the older generation it does not seem to be the case that the presence of siblings is terribly important in explaining successful marriages and divorce. It is worth noting that in the case of the CHARLS data set the vast majority of respondents are over age 45 and as such, very few of them are only children. If we restrict the sample to just those that are only children (column 9), we find that the

sample of ever married only children shrinks to less than 400.

We next turn out attention to the CFPS data and see how this compares to the results suggested by the CHARLS data. Table 3.4 presents regression results. We find that the results are somewhat similar to that of the CHARLS analysis in that the coefficient is quite small and significant in some specifications, but only at the 90% confidence interval. In terms of magnitude, the presence of a sibling is associated with a 1‰decline in the likelihood of divorce. Thus, the result is nearly identical to our estimate using CHARLS data. However, there are a great many more only children in the CFPS data so we feel relatively more confident in this result. The relationship appears slightly more pronounced for the relatively highly educated as shown in column 4 of Table 3.4. However, we again only find this to be significant at the 90% confidence interval. Although the presence of siblings is associated with a 3‰decline in the likelihood of divorce, this is certainly not supportive of the popular interpretation that only children's failure to maintain successful marriages is a primary driver of divorce in China.

Given the very small magnitudes of our estimates across specifications and data samples, we conclude that the increasing prevalence of only children in China is not a primary driver of the recent increases in divorce. Moreover, it would appear that the presence of siblings is associated with a very similar decrease in the probability of divorce for both the older generation of Chinese and a younger one as well.

3.5.3   Role of proxies for Western Attitudes

Use of a Lunar Calendar Using CHARLS Data

Another common explanation for the increased prevalence of divorce in China is that many traditional attitudes about marriage have been replaced by Western ideals that are more accommodating of divorce. Unfortunately, the presence and acceptance of Western attitudes is inherently unobservable. Despite this, there are some interesting proxies that are observed in the data. One such example is that in China individuals often differ in how they report their dates of birth. Roughly two thirds of the sample uses the same western calendar as is used for day to day activities. However, another segment of the population instead reports their date of birth using the traditional lunar calendar that is far more closely tied to traditional ceremonies and spirituality. Whether one chooses to use the Western calendar or the lunar calendar would appear to be an excellent proxy for the extent to which one ascribes to more traditional Chinese attitudes about marriage and divorce.

We test this possibility by using CHARLS data as this survey allows respondents to report dates of birth and death using the lunar calendar. Interestingly, the answer seems to be that use of the lunar calendar is significant in explaining divorce tendencies. Table 3.1 suggests that the relationship between divorce and the lunar calendar is significant at the 99 percent confidence interval and the sign of the coefficient is negative. Thus, those that use the lunar calendar are less likely to get a divorce. In terms of a magnitude across specifications it appears that most estimates are just slightly over 0.01 thus, using the lunar calendar is associated with roughly a 1% decline in the probability of divorce. This would

lend support to the claim that traditionally minded people, who are more likely to use a lunar calendar, are more likely to view marriage as being permanent.

A related, but surprising result is that village dwellers are more likely to become divorced than their urban counterparts. In particular, Table 3.1 suggests village residence is associated with a 3% increase in the likelihood of divorce and is significant at the 99% confidence interval. Although this appears somewhat surprising, it is possible that the dummy for lunar calendar is accounting for traditional thinking about marriage. In a regression specification without the lunar calendar dummy, village residence is associated with lower divorce. However, once we control for traditional mindsets by using a lunar calendar indicator, the sign switches and becomes positive. This suggest that a less nuanced investigation might erroneously conclude that rural people are less likely to divorce, when the decrease is a function of traditional attitudes, and not rural living.

Another consideration is to see if lunar calendar use remains important for city dwellers, and within the sample of village dwellers. Hence, we divide the sample and estimate the previous regression for these two distinct samples. This is shown in Table 3.1, columns 4 and 5. If we restrict the sample to just village residents, we again find that use of the lunar calendar is significant. The result is robust in the urban case as well. A potential interpretation is that regardless of location, more traditionally minded individuals are more likely to use the lunar calendar, and are also less likely to divorce. Similarly, these individuals are less likely to have adopted Western attitudes about divorce and are less likely to divorce. This supports the interpretation that within the group of village or city dwellers there are individuals that are more traditional minded, and importantly, those

more traditional minded individuals are less likely to divorce.

Age at First Marriage

Traditionally, Chinese are expected to marry fairly young by current Western standards. However, since the unification of China under the Communist Party the typical age at marriage has risen and now the mean ages for first marriage are roughly 23 for women and 25 for men. However, there remains considerable heterogeneity in age at first marriage. In particular, a young age at first marriage may be considered a good proxy for a traditional mindset. We seek to examine whether or not age at first marriage is predictive of divorce In general, findings are supportive of the interpretation that later marriage is associated with a higher likelihood of divorce, moreover the relationship is highly significant across specifications. Table 3.1 shows that as the age of first marriage increases by one year, the likelihood of divorce rises by roughly 2‰. Although this may appear very modest, it is important to note that this is a per year estimate. That is for every year that marriage is delayed, the probability of divorce rises by 2‰. Thus, we do not find the implication uninteresting.

This may appear at first glance to be surprising as it is rather different than what we might expect in modern Western nations (specifically we tend to think of people who marry very young as being at a greater risk of divorce). Moreover, from a purely quantitative perspective, we might think that as years married increases the cumulative probability of divorce increases. However, in the case of the Chinese, younger ages at marriage are very likely associated with more traditional attitudes and expectations of marriage. It is likely for this reason, that we see such a robust relationship between age at first marriage and

127

incidence of divorce. It is worth noting that this is not a consequence of omitting observable variables positively associated with age at first marriage such as education, village or urban dwelling, or the other proxy for western thinking - use of the lunar calendar. Thus, the relationship between divorce and age at 1st marriage, is significant independent of these observable and included covariates.

We also consider the age at first marriage using the CFPS data. Table 3.4 shows results that are rather different than for the CHARLS data. In the case of the CFPS data, we again find that later marriage is associated with a greater degree of divorce, but the result is not distinguishable from zero. Moreover, the sample size is generally quite large and estimates are reasonably precise. This would suggest that for the older generation of Chinese, early marriage was a good proxy for traditional attitudes and an indicative of a lower probability of divorce. However, given results using CFPS data it would seem that this is not a useful predictor of divorce for a younger generation of Chinese. This may simply indicate that age at marriage is no longer important, or that age at marriage is merely a proxy for traditional attitudes which do matter. However, it is only a reliable proxy for older Chinese and not younger Chinese.

Importance of Mandarin using CFPS

Although the CFPS doesn't allow for respondents to provide important dates using the lunar calendar, there is another question that may be deeply revealing of the degree to which individuals are traditional minded or not. In particular, the CFSP asks respondents how important the feel an intimate knowledge of the Mandarin Chinese language is. Our thinking is that in order to be connected with media, education, non-agricultural job op-

portunities, and a life outside of the ancestral village, knowledge of Mandarin is absolutely necessary. The alternative is to only know one's very specific dialect, which will still allow one to engage with their immediate locality, but not much else. Since traditional minded individuals are generally much less interested in life beyond their immediate sphere, they will likely not value Mandarin as much as more modern minded individuals. On the other hand, those with an outward disposition and a desire to engage with the modern world will feel that Mandarin is absolutely necessary.

A point to be made is that we include a measure of Mandarin importance only after controlling for age, province and educational attainment. Thus, the measure will not be confounded with provincially based language preferences, the general trend towards greater Mandarin acceptance over time, and the fact the formal education is conducted in Mandarin. We find that attitudes about the relative importance of Mandarin are associated with the incidence of divorce. People that feel Mandarin is relatively more important are more associated with divorce. This reflects our interpretation that those that view Mandarin as important, are less traditional minded, and therefore more likely to divorce. Table 3.4 shows our estimates for the relationship between Mandarin importance and divorce. Our result is significant, but only for our preferred specification (column 1). Once we divide the sample into different segments of the population (columns 3-6), standard errors grow and the result loses any significance. Thus, the result is an interesting one, but not a robust one.

3.5.4 Economic Explanations

A variety of economic explanations are worth considering in terms of explaining Chinese divorce. Among them are the effects of education, economic migration, income and relative affluence.

Education

Education may be important for at least two prominent reasons. One is that education may alter an individual's beliefs and expectations about what constitutes an appropriate marriage. Available data do not allow us to consider how education may affect beliefs and expectations, but we are able to see how education has a bearing on divorce. The other fairly salient possibility is that education alters the opportunity cost of marriage. Quite simply, the idea is that marriage may require economic sacrifices that are likely to be more dramatic for the relatively well educated, than for those with more modest levels of education. This is a common explanation for the recent surge in divorces, essentially the argument suggests that marriage constrains economic mobility and the ability to pursue a variety of economic opportunities. As individuals become more educated the value of these lost opportunities rises. If this is the case, it is very likely that higher levels of educational attainment will be associated with a greater incidence of divorce.

Another related interpretation is that as individuals gain education they will find that the gains from intra-household specialization declines precipitously and divorce becomes a relatively more appealing option. Of course, the flip side of the same coin is that educational increases may lead to less assortative matching. As a result, marital partners

are more likely to have similar educational backgrounds and consumption complementari-ties may rise, making marriage more attractive. Separating these two possibilities is beyond the scope of this paper, but we hope to indicate how and if education is associated with divorce.

We consider educational attainment and how it relates to divorce. The CHARLS survey is somewhat limited in that it only asks if respondents completed primary school or middle school. It does not further refine its questions to ask about high school or post high school education. Although this may seem bizarre, it is important to recognize that the sample in the CHARLS dataset is primarily those aged 45 and up. Thus, at the time these individuals received their education, China was still a very poor and developing nation where illiteracy was quite high. Moreover, the dataset includes those that would have received some additional education, but did not as the Cultural Revolution rendered many forms of further education politically problematic.

Table 3.1 shows the results of our estimation. We use dummies for primary school attainment, and middle or higher attainment. The omitted or baseline group are those that did not complete primary school. This is roughly 14% of the surveyed population. We find that there is not a significant relationship between education and divorce. This would seem to lend support to the interpretation that at least for older generations, higher educational attainment is not associated with more divorce. It remains possible that by altering expectations and opportunity costs, education does impact divorce, but in opposite ways of a similar magnitude. Thus, lost gains from intra-household specialization push divorces up, while gains from consumption complementarities diminish divorce by a similar

magnitude. Although we cannot rule this out, it strikes us as rather unlikely. A potential problem in the results is that the data set does not identify highly educated individuals, thus we are not able to determine if high levels of education are associated with a greater or lesser degree of divorce.

We also consider the effect of education for respondents to the CFPS survey. This may potentially be of greater interest for two key reasons. One is that educational attainment has risen in recent years, therefore the CHARLS data and its older respondents are potentially not representative of those entering into marriage today. The other reason, is that CFPS is far more specific than is the CHARLS survey in asking people about their educational attainment. CFPS allows researchers to identify schooling to within half a year. We consider a variety of ways that education may enter into our regression, but the central finding remains the same. That is, education does not appear to be associated with a change in the probability of divorce. Table 3.4 allows years of education to enter into the regression on a year by year incremental basis. Across specifications and populations the result is not distinguishable from zero. We also allowed for education to enter with polynomial terms and used dummies for discrete educational cutoffs in the event that education may interact non-linearly with divorce. In the case of these alternative methods, results were the same, but with less precision and lower R^2.

This finding would seem to suggest that education is not terribly important in explaining divorce even after correcting for some of the shortcomings of the CHARLS data.

Migration

Migration within China may also be an issue worth devoting some attention to. Ostensibly, the idea being that those that migrate are likely to place a greater value on economic opportunities. As such, they may feel the economic opportunity costs of marriage more acutely. Of course this decision is likely not reached in a vacuum, rather there may be some degree of endogeneity in the decision process. That is, couple who feel their marriage is particularly strong may be less deterred by the prospect of migration. If it is the case that migrants value economic opportunity more highly and do not because they systematically feel more stable in their marriages, we would expect to find that migrants are generally more likely to experience divorce. We find this is not the case as shown in Table 3.1. Across specifications and populations, it appears being a migrant is not associated with greater or lesser incidence of divorce.

This may be universally true, or may only be true for the elderly. It is quite possible that for the older generation migration was not particularly associated with the pursuit of greater economic opportunity. However, today migration is inexorably linked to labor market activities as most individuals must sacrifice their access to publicly provided goods and services if they migrate. A peculiarity of Chinese domestic policy is that although all citizens are able to receive certain social services, they must be residing in their approved locality (usually the place of their birth) in order to collect services and benefits. As such, today people generally only migrate if they believe the benefits will outweigh the costs of lost access to social services, such as education and health care provision.

We also consider migration using the CFPS data set. Table 3.4 shows that results

are generally very different for this data than for CHARLS data. In particular, not having migrated is associated with a diminished likelihood of divorce. The result is significant at the 99th percent confidence interval, and is robust to changes in specification and changes in the population considered. The magnitude suggest not having migrated decreases the likelihood of divorce by between 2% and 3%. As to why this is so different than in the CHARLS dataset, our thinking is that migration was less likely to be for economic reasons for the older generation, as such it likely did not radically alter the opportunity costs of marriage. This stems from the fact that levels of economic development across China were far less heterogenous prior to the post Tiananmen economic reforms. As China has undergone rapid development, the economic value of migration has likely increased a great deal, as some parts of China became far wealthier and were able to offer desirable employment to many people.

In terms of interpreting the result within a marriage framework, two thoughts come to mind. Migration may result in the couple living apart and/or having to adapt to a stressful new environment. This is likely not good for marriage prospects. Our second thought, is that in terms of a more traditional economic model, couples that migrate are probably those that are more career oriented and more likely to feel the loss of economic opportunities due to marriage. Furthermore, migration is more likely to result in dual earner households where gains from intra-household specialization are diminished. This would also likely increase divorce in a *ceteris paribus* setting.

Income

Income disparities across married couples may also be important in explaining divorce. On the one hand, higher incomes may be associated with those that value career relatively more, those with higher levels of education, and those with more modern attitudes about marriage and/or divorce. This would suggest divorce rates will be higher for the affluent. On the other hand, life may be relatively less stressful for the affluent and consumption complementarities may be more pronounced for households with greater consumption. This would suggest, divorce is less likely for those with higher incomes.

The CHARLS data set does provide income information, but it is limited to income in the previous month, which is quite noisy with many individuals reporting an income of zero. Moreover, there are a great many observations which are missing values, as well as some observations that appear highly suspect (reports of incomes much greater than would be normally expected). As such, it appears this particular independent variable is at best very noisy, and at worst, unreliable. Given this, we chose not to make use of income as an independent variable. Instead, we have used a measure of relative affluence.

To obtain a measure of self perceived relative affluence, surveyors present the respondent with a fictitious individual referred to as Mr. Wang. The survey participant is told a good deal about Mr. Wang's income, assets and leisure time. The respondent is then asked to evaluate how wealthy they are relative to the fictitious Mr. Wang. Responses are restricted to: much worse than Mr. Wang, slightly worse than Mr. Wang, about the same as Mr. Wang, slightly better than Mr. Wang, and much better than Mr. Wang. The respective frequencies of each of these 5 responses is roughly 10%, 26%, 29%, 23%, 7%.

According to CHARLS administrators, Mr. Wang is intended to be viewed as fairly middle class. We use this as an independent variable to account for how wealthy the respondent feels from a relative perspective. Findings are presented in Table 3.1. The variable name is SR Wealth for self-reported relative wealth. In short, it seems those that view Mr. Wang as worse off then themselves are quite a bit more likely to have divorced at some point in their life. The result is highly significant and robust to changes in specification. Roughly speaking, viewing oneself as one increment wealthier (among the five options) is associated with a 1 to 2 percent increase in the likelihood of divorce. This suggests that self perceived economic success may be associated with a greater incidence of divorce in China. The relationship is very significant and distinguishable from zero.

It is difficult to extrapolate from this how income alters divorce, since our measure is both relative and self-reported, but is still a result of some interest. It lends support to the idea that as China has become wealthier, divorces have risen. However, it is important to note that this measure is entirely about relative wealth and not absolute wealth. As such, it may be more correct to think of China as experiencing more divorce as the society becomes economically less equal. Finally, as with all results from the CHARLS data, it is imperative to keep in mind that the result is for those age 45 and older at the time of the survey. Thus, the result may not hold for today's generation.

CFPS data presents income in a much more expected fashion having measures for income at the individual and household level for the past year and month. Moreover, the data do not appear to present any discomfiting irregularities. Additionally, there are no questions such as the Mr. Wang question in the CHARLS data. Thus, we consider inclusion

of income into our regression framework from a variety of perspectives. We estimate the coefficient value using multiple measures of income, but we ultimately settle upon using the log of annual income as this is the standard given the fact that incomes can become quite far from the center of the distribution, and also because of the diminishing marginal value of income.

None of the measures suggest a relationship distinguishable from zero. We find that the log of income does not suggest the presence of a relationship between income and divorce. This suggests that high income households and low income households do not face radically different divorce probabilities. In terms of reconciling this result with CHARLS data results, we recognize that this is not straightforward and an interpretation is very likely to be incorrect. Having said that, we feel the two results may not be truly comparable since one is a measure of absolute income, while the other is a measure of relative well being from a wealth and leisure perspective.

3.5.5   Health and Divorce

The CHARLS dataset is quite well suited to considering the long run health implications of divorce, in that the sample is mostly comprised of middle-aged and elderly respondents, including many that have experienced a divorce. We examine their physical and mental health to see if those that have experienced a divorce are generally different in their health outcomes, than those who have not divorced. A potentially surprising trend in China is that the majority of the ever divorced population is able to remarry at a later date in time. As such, a significant subset of the ever-divorced population has re-married. Therefore, it is of interest to consider how their heath outcomes differ, if at all, from the married

and never divorced, and the divorced but not re-married individuals. A final consideration is whether adverse health outcomes are purely mental, physical, or both.

In theory, a divorce is a mentally and emotionally taxing experience that may result in worse health outcomes. This may be true from both a physical and mental health perspective. Moreover, prior research by Amato (2000) suggests this is a common occurrence for Westerners who divorce. Moreover, Amato's research suggests these effects may still be present for the remarried, but to a lesser extent than for those that divorce and remarry. We wish to be unambiguously clear that our analysis is descriptive, and not causal. A key criticism of studies linking divorce with health outcomes astutely notes that there may be significant reverse causation. That is, divorce may be a result of poor health, and not simply an outcome. Moreover, the population of individuals who divorce are likely to be different than those that do not experience divorce in a great many observable and unobservable ways. Failure to account for these differences, which is virtually guaranteed in an observational setting, will result in biased coefficients. Our analysis, like others that have preceded it, is clearly not immune to these problems.

Despite identification short comings, we wish to see to what extent Chinese data is supportive of Western findings that poor health is associated with divorce. To consider this we regress a host of self reported health outcomes on the following specification, and explicitly allow for dummies that indicate whether a person is remarried, or divorced without having remarried.

$$health\_measure = \alpha + \beta currently\_divorced + \beta remarried + \gamma controls + \epsilon$$

There are some individuals that have been married and divorced on multiple occasions, as these individuals are not common, nor easily categorized, they are omitted from consideration. The omitted reference group is the population that has married, but never divorced. Widowed and never married individuals are dropped. Findings are reported in Table 3.2. In general, findings suggest that divorce is associated with worse self-reported health. The act of remarrying appears to diminish some of these adverse effects, but is unable to reduce them to the levels found in the married, never divorced sample.

Mental Health In general, findings are supportive of the claim that divorce is harmful to mental health and remarriage appears to alleviate some of the mental burden, but not all of it. For example, Table 3.3, columns 5 through 10, show that the divorced, but un-remarried, are more likely to report feeling lonely, not wanting to continue living and report lower levels of life satisfaction. In all cases, the result is highly significant and distinguishable from zero. However, in each of these cases the magnitude of the coefficient is considerable smaller for the remarried portion of the sample. This would suggest remarrying is at least somewhat effective in reducing some of the mental trauma associated with divorce. However, it is worth noting that the remarried still exhibit worse outcomes relative to the never divorced portion of the sample, and the relationship for remarried individuals is also very significant and distinguishable from zero.

In the case of feeling fearful about the future, the relationship appears to be that being divorced has no effect, but being re-married is associated with an elevated probability of feeling fearful about the future. Columns 3 and 4 of Table 3.3 suggest this pattern in that the coefficient for the divorced population is not distinguishable from zero. However,

it appears that the remarried are nearly 10% more likely to feel fearful about the future than the never divorced, and the relationship is very significant. Notably, the magnitude of the coefficient is higher for the divorced, but less precise. However, this odd outcome is most likely explained by the relatively fewer observations where individuals divorce, but then do not remarry. As such, the coefficient estimate is inherently less precise and for this reason we are unable to unequivocally point to the existence of a relationship between feeling fearful, and having been divorced, but not remarried.

The relationship is essentially the same for feeling hopeful about the future. Columns 1 and 2 in Table 3.3 reveal the relationship between divorce, remarriage and feeling hopeful about the future. In short, we are unable to rigorously point to a relationship between divorce and feeling hopeful about the future, but we can make the claim that there is a highly significant relationship for the remarried, who are less likely to report feeling hopeful about the future. Again, the magnitude of the coefficient is larger for the remarried, but by virtue of fewer observations the coefficient estimate is much less precise.

The CFPS does, however, have several measures of mental health, much like the CHARLS, but they are slightly different as outlined above. These may be slightly less problematic than physical health, since it is less clear that these particular indicators for mental health are fairly rare for younger people. Thus, the fact that CFPS has relatively fewer elderly respondents, may be less problematic from the perspective of mental health, than from the perspective of physical health.

Table 3.5, columns 2-6 show our coefficient estimates for dependent variables related to mental health. In general, results are supportive of what was gleaned using

140

CHARLS data. Namely, mental health outcomes are worse for both the divorced and un-remarried, and the remarried, than it is for those who married, but never divorced. In terms of a ranking the best off are those that marry and never divorce, followed by the divorced and remarried, followed by the divorced, but not remarried who fare the worst. Within the context of the table, the baseline comparison group is composed of individuals that have married but never divorced.

Being divorced is associated with a worse outcome for all considered measures, and the findings are significant at 99% confidence intervals. That is the divorced and un-remarried are more likely to score highly on measures of depression, pessimism. Yet, they are more likely to score lowly on measures of social satisfaction, optimism and happiness. However, for the remarried group, only some findings are significant. Namely, the remarried are less likely to feel happy or socially satisfied at 99% confidence intervals. In both cases the magnitude is quite a bit less than for the divorced, un-remarried population. Furthermore, the remarried are slightly less likely to score highly on a measure of optimism, but the result is only significant at the 90% confidence interval.

Physical Health The relationship between, divorce, remarriage and physical health out-comes is generally less easily characterized than the relationship with mental health. Two variables of interest the CHARLS allows us to consider are the prevalence of chronic dis-ease, and average hours of sleep per night. Chronic diseases include diseases such as cancer, diabetes, cardiovascular diseases, and other long term non-communicable ailments. Results are shown in Table 3.2.

The relationship between divorce and chronic disease matches expectations fairly

well, in that relative to the married but never divorced population, the remarried population is significantly more likely to suffer from a chronic disease. The thinking behind this result is that divorce is a traumatic experience they may have a bearing on physical health. Remarriage alleviates some of these problems, but is unable to completely "undo" some of the adverse outcomes associated with divorce. This is mirrored in research by Amato (2000) using western data sources. Indeed, in the CHARLS sample the remarried, are 5% more likely to have a chronic disease, and the result is very significant. In the case of the divorced, but not remarried population the effect is not distinguishable from zero. This appears to be driven by imprecision in estimating the coefficient, and likely stems from the fairly small sample size of divorced, but un-remarried individuals.

The relationship between hours of sleep and previous divorce is less easily characterized and unexpected. Table 3.2, columns (3) and (4) show that although the divorced and un-remarried are generally unable to sleep as much as other segments of the population, the relationship is again fairly imprecise. More unexpectedly, it seems that the remarried actually sleep more hours than the married, but never divorced population. Moreover, this relationship is very distinguishable from zero. The finding that the divorced and remarried typically sleep more than married, never divorced people is quite surprising.

We also make use of the CFPS to consider how divorce may impact health. The CFPS is perhaps less well suited than the CHARLS data for one clear reason. That is that the CHARLS data is almost entirely those old enough to begin experiencing chronic illness, while the CFPS contains a great many individuals that are likely to young in most instances to have first hand experience with chronic disease.

We estimate the following equation:

$$health\_measure = \alpha + \beta_c urrently\_divorced + \beta remarried + \gamma controls + \epsilon$$

For dependent variables we consider measures of: Presence of a Chronic Disease, Depression, Feeling Pessimistic about the Future, Self-reported Overall Happiness, Degree of Social Satisfaction and Feeling Optimistic about the Future. All measures are self reported. Chronic disease is binary. Depression, happiness, and social satisfaction are scaled on a measure from 1 to 30 based on responses to other questions that we do not have access to. Feeling Pessimistic or optimistic about the future are continuous measures from zero to 1 and again are obtained by virtue of questions we do not have access to.

Table 3.3: Divorce and Related Mental Health Outcomes

VARIABLES	(1) Hopeful	(2) Hopeful	(3) Lonely	(4) Lonely	(5) Continue Life	(6) Continue Life	(7) Overall Satisfaction	(8) Overall Satisfaction
Remain Divorced	-11.3		55.8***		18.7***		20.4***	
	(10.6)		(8.62)		(6.91)		(6.22)	
Remarried		-16.6***		21.9***		9.15**		6.97*
		(6.09)		(4.98)		(3.97)		(3.64)
Age	-1.37***	-1.35***	0.649***	0.613***	0.433***	0.419***	-0.578***	-0.592***
	(0.130)	(0.130)	(0.106)	(0.106)	(0.0845)	(0.0846)	(0.0780)	(0.0781)
Lunar	-0.147***	-0.148***	0.0602**	0.0579**	0.0638***	0.0633***	0.0226	0.0216
	(3.12)	(3.12)	(2.56)	(2.56)	(2.03)	(2.03)	(1.83)	(1.83)
From Village	4.64	4.49	-6.04*	-4.41	-6.28**	-5.75**	0.728	0.732
	(4.30)	(4.29)	(3.53)	(3.53)	(2.81)	(2.80)	(2.52)	(2.51)
Educ Primary	12.2***	12.2***	-13.0***	-12.9***	-11.0***	-10.9***	0.252	0.288
	(3.21)	(3.21)	(2.63)	(2.63)	(2.09)	(2.09)	(1.90)	(1.90)
Educ Middle	12.7***	12.7***	-10.4***	-10.2***	-7.96***	-7.91***	-2.75	-2.70
	(3.41)	(3.41)	(2.80)	(2.81)	(2.23)	(2.23)	(2.00)	(2.00)
Sibling	0.532	0.527	-0.596	-0.626	-0.0961	-0.104	-0.180	-0.202
	(0.643)	(0.642)	(0.524)	(0.525)	(0.417)	(0.417)	(0.379)	(0.379)
Age First Marriage	0.178	0.113	-1.05***	-0.923***	-0.874***	-0.826***	0.483***	0.526***
	(0.310)	(0.310)	(0.252)	(0.253)	(0.200)	(0.200)	(0.183)	(0.183)
SR Wealth	-21.2***	-21.1***	20.8***	20.9***	22.2***	22.2***	30.0***	30.0***
	(1.62)	(1.62)	(1.33)	(1.33)	(1.06)	(1.06)	(0.959)	(0.960)
Constant	437.8***	438.6***	76.7***	75.1***	54.6***	53.9***	208.8***	208.3***
	(13.8)	(13.8)	(11.2)	(11.2)	(8.94)	(8.94)	(8.16)	(8.17)
Observations	9,111	9,111	9,256	9,256	9,243	9,243	8,434	8,434
R-squared	0.156	0.156	0.159	0.157	0.177	0.177	0.217	0.217
Data Source	CHARLS	CHARLS	CHARLS	CHARLS	CHARLS	CHARLS	CHARLS	CHARLS

SE in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Coefficients and SE are multiplied by 100 for ease of interpretation

Table 3.4: Predictors of Divorce

Specification Dependent variable	(1) Total population divorce	(2) Total Population divorce	(3) Low Educ divorce	(4) High Educ divorce	(5) With Siblings divorce	(6) Only Child divorce
Age	-0.008.49	-0.00990	-0.0230**	0.0764**	-0.0126	-0.00270
	(0.00989)	(0.0102)	(0.0107)	(0.0313)	(0.0109)	(0.0332)
Female	0.979***	1.03***	1.32***	-0.0465	1.13***	-0.0771
	(0.263)	(0.264)	(0.284)	(0.690)	(0.276)	(0.972)
Urban	1.67***	1.33***	1.17***	2.47***	1.41***	0.854
	(0.272)	(0.284)	(0.295)	(0.866)	(0.294)	(1.17)
No Migration	-2.78***	-2.45***	-1.38**	-3.93***	-2.36***	-3.40*
	(0.528)	(0.533)	(0.636)	(1.05)	(0.560)	(1.75)
Siblings	-0.113*	-0.103	-0.0824	-0.358*		
	(0.0671)	(0.0680)	(0.0709)	(0.204)		
Mandarin Importance	0.280**	0.164	0.162	0.144	0.151	0.364
	(0.112)	(0.114)	(0.115)	(0.400)	(0.118)	(0.473)
Log Income	-0.0537	-0.0605	-0.0199	-0.318	-0.0267	-0.657**
	(0.0880)	(0.0899)	(0.0929)	(0.278)	(0.0940)	(0.333)
Age First Marriage	0.0183	0.00904	-0.0275	0.262***	-0.00601	0.203*
	(0.0322)	(0.0327)	(0.0338)	(0.0991)	(0.0341)	(0.116)
Years of Educ	0.0477	0.0344	0.00836	-0.296	0.0190	0.241*
	(0.0308)	(0.0313)	(0.0393)	(0.225)	(0.0324)	(0.125)
Constant	4.61***	8.68***	8.53**	19.4***	7.74***	19.3***
	(1.32)	(2.49)	(3.81)	(5.55)	(2.83)	(6.40)
Observations	20,570	20,503	16,333	4,170	18,956	1,547
Geographic Controls	No	Yes	Yes	Yes	Yes	Yes
Ethinicity Controls	No	Yes	Yes	Yes	Yes	Yes
R-squared	0.007	0.017	0.017	0.035	0.018	0.050
Data Source	CFPS	CFPS	CFPS	CFPS	CFPS	CFPS

SE in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Coefficients and SE are multiplied by 100 for ease of interpretation

Table 3.5: Divorce and Related Physical and Mental Health Outcomes

Dependent Variable	(1) Chronic Disease	(2) Depression	(3) Pessimistic	(4) Happiness	(5) Social Satisfaction	(6) Optimistic
Remain Divorced	2.07	-103.0***	-26.9***	-80.9***	-12.5***	-29.7***
	(1.96)	(20.7)	(5.35)	(5.48)	(4.60)	(6.10)
Remarried	2.86	-28.6	-7.43	-19.5***	-15.6***	-11.9*
	(02.00)	(21.1)	(5.44)	(5.59)	(4.69)	(6.22)
Age	0.537***	-0.517**	-0.151***	0.286***	0.525***	-1.37***
	(0.0207)	(0.219)	(0.0565)	(0.0577)	(0.0485)	(0.0643)
Female	-3.21***	44.2***	11.4***	-5.53***	-6.83***	6.19***
	(0.535)	(5.67)	(1.46)	(1.49)	(1.26)	(1.66)
Urban	-0.205	13.5**	3.36**	5.53***	0.0787	-7.38***
	(0.575)	(6.09)	(1.57)	(1.61)	(1.35)	(1.79)
No Migration	-3.70***	33.1***	8.40***	2.00	2.21	17.6***
	(1.08)	(11.4)	(2.94)	(3.01)	(2.53)	(3.35)
Siblings	0.675***	-5.41***	-1.41***	-2.85***	-0.138	-0.980**
	(0.137)	(1.46)	(0.376)	(0.384)	(0.323)	(0.428)
Mandarin Importance	0.863***	9.39***	2.48***	10.2***	12.3***	10.8***
	(0.231)	(2.45)	(0.632)	(0.644)	(0.542)	(0.718)
Age First Marriage	0.161**	1.62**	0.419**	1.06***	1.04***	0.882***
	(0.0662)	(0.701)	(0.181)	(0.185)	(0.155)	(0.206)
Constant	-11.8**	2567.0***	-31.5**	327.7***	328.7***	329.8***
	(5.04)	(53.2)	(13.7)	(14.1)	(11.8)	(15.8)
Observations	20,491	20,333	20,333	20,489	20,477	20,451
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethinicity Controls	Yes	Yes	Yes	Yes	Yes	Yes
Income and Educ Controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.058	0.064	0.064	0.076	0.068	0.079
Data Source	CFPS	CFPS	CFPS	CFPS	CFPS	CFPS

SE in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Coefficients and SE are multiplied by 100 for ease of interpretation

A shortcoming of the CFPS data is that there is only one measure of physical health present, namely whether or not the respondent suffers from a chronic illness. The respondent is not restricted to some subset of chronic disease, but rather is given an explanation for what constitutes a chronic disease and is then asked if they feel that they have one or not. Table 3.5, column 1 suggests that the divorced and remarried are not different than the married, but never divorced population. In the case of the divorced and remarried, they both have positive coefficients, but the estimates are not significant at any level. Thus, they are not distinguishable from zero. This would suggest that results from the CHARLS are either not robust to changes in data source, or that the CFPS does not contain a sufficient number of people that are of an age where they are likely to experience chronic illness.

3.6 Conclusion

Given the recent dramatic increase in divorce in mainland China, a plethora of explanations have been circulated in the media and public discourse. Moreover, the Economics literature has also provided a variety of explanations and associated issues to consider. In light of these considerations, we consider which factors are associated with divorce in China using two large, detailed, national samples that reveal individuals' marriage histories. One sample is more representative of the over 45 population, and the other sample is more representative of the nation as a whole. In doing so, we hope to see how marriage and divorce have changed in China. We find that Western or modern attitudes, as measured by several proxies, are quite predictive of divorce, moreover individuals that feel they are relatively well off compared to others are also more likely to divorce. Interestingly, many attributes

one might expect to be predictive of divorce, are not. These fairly unimportant predictors include, education, income and age at marriage. The relationship between migration and divorce depends on the sample. Since the result varies by data source, we do not wish to make a strong claim, but note that migration is a strong predictor for the relatively younger set of respondents in the CFPS, but migration is not meaningful for the relatively older population surveyed in the CHARLS.

We also consider how divorce may affect physical and mental health. We find that both surveys lend strong support to the interpretation of divorce as being associated with worse mental health outcomes. However, we find mixed evidence that divorce is associated with worse physical health outcomes. Namely, we find that the remarried fare worse than the married, never divorced population. However for the divorced and un-remarried population, estimates are less precise as the sample size is much smaller. Thus, we are unable to marshall significant evidence that physical outcomes are worse for the divorced, but unremarried population.

Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010), "Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program." *Journal of the American Statistical Association.*

Allen, Douglas (1992), "Marriage and divorce: Comment." *American Economic Review.*

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach (2008), "Inside the war on poverty: The impact of food stamps on birth outcomes." *National Bureau of Economic Research Working Paper 14306.*

Amato, Paul R. (2000), "The consequences of divorce for adults and children." *Journal of Marriage and Family.*

Andrews, Margaret and Parke E. Wilde (2000), "The food stamp program in an era of welfare reform: Electronic benefits and changing sources of cash income." *Journal of Consumer Affairs*, 34.

Astrup, Poul, H.M. Olsen, Dyre Trolle, and Knud Kjeldsen (1972), "Effect of moderate carbon-monoxide exposure on fetal development." *The Lancet.*

Bailey, Martha (2010), "Momma's got the pill: How anthony comstock and griswold v. connecticut shaped U.S. childbearing." *American Economic Review.*

Bastiotis, Peter, Carol S. Cramer-LeBlanc, and Eileen T. Kennedy (1998), "Maintaining nutritional security and diet quality: The role of the food stamp program and wic." *Family Economics and Nutritional Review.*

Becker, Gary S. (1981), *A Treatise on the Family.* Harvard University Press, Cambridge, MA.

Behera, Dinesh, Salim Dash, and Salman Malik (1988), "Blood carboxyhaemoglobin levels following acute exposure to smoke of biomass fuel." *Indian Journal of Medical Results.*

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004a), "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics.*

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004b), "How much should we trust differences-in-differences estimates?" *Quarterly Journal of Economics.*

Blau, Francine D. and Lawrence M. Kahn (2000), "Gender differences in pay." *Journal of Economic Perspectives.*

Blundell, Richard and Luigi Pistaferri (2003), "Income volatility and household consumption: The impact of food assistance programs." *Journal of Human Resources.*

Boy, Erick, Nigel Bruce, and Hector Delgado (2002), "Birth weight and exposure to kitchen wood smoke during pregnancy in rural guatemala." *Environmental Health Perspectives.*

Butler, J.S. and J.E. Raymond (1996), "The effect of the food stamp program on nutrient intake." *Economic Inquiry.*

Card, David and Alan B. Krueger (1994), "Minimum wage and employment: A case study of the fast-food industry in new jersey and pennsylvania." *The American Economic Review.*

Chay, Kenneth and Michael Greenstone (2003a), "Air quality, infant mortality, and the clean air act of 1970." *NBER Working Paper.*

Chay, Kenneth and Michael Greenstone (2003b), "The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession." *Quarterly Journal of Economics.*

Chen, Long, Yang Wei, Brian Jennison, and Stanley Omaye (2002), "Air pollution and birth weight in northern nevada." *Inhalation Toxicology.*

Chen, Yuyu, Avraham Eberstein, Michael Greenstone, and Hongbin Li (2014), "Evidence on the impact of sustained exposure to air pollution on life expectancy from china's huai river policy." *Proceedings of the National Academy of Sciences.*

Currie, Janet (2003), *U.S. Food and Nutrition Programs.* University of Chicago Press, Chicago.

Danielson, Caroline and Jacob Alex Klerman (2006), "Why did the food stamp caseload decline and rise? effects of policies on the economy." *Rand Corporation Working paper, WR386.*

Dary, Omar, Oscar Pineda, and Jose Maria Belizan (1981), "Carbon monoxide contamina-

tion in dwellings in poor rural areas of guatemala." *Bulletin of Environmental Contamination and Toxicology.*

David Cutler, Edward Glaeser and Jesse Shapiro (2003), "Why have americans become more obese." *Journal of Economic Perspectives.*

Diamant, Niel J. (2000), *Revolutionizing the Family: Politics, Love and Divorce in Urban and Rural China.* University of California Press.

Donald, Stephen G. and Kevin Lang (2007), "Inference with difference in differences and other panel data." *The Review of Economics and Statistics.*

Elizabeth Caucutt, Nezih Guner and John Knowles (2002), "Why do women wait? matching, wage inequality, and the incentives for fertility delay." *Review of Economic Dynamics.*

Fraker, Thomas and Robert Moffitt (1998), "The effect of food stamps on labor supply: A bivariate selection model." *Journal of Public Economics.*

Friedberg, Leora (1998), "Did unilateral divorce raise divorce rates? evidence from panel data." *American Economic Review.*

Gary S. Becker, Elizabeth M. Landes and Robert T. Michael (1977), "An economic analysis of marital instability." *Journal of Political Economy.*

George Akerlof, Janet Yellen and Michael Katz (1996), "An analysis of out-of-wedlock childbearing in the U.S." *Quarterly Journal of Economics.*

Goldin, Claudia and Lawrence Katz (2002), "The power of the pill: Oral contraceptives and women's career and marriage decisions." *Journal of Political Economy.*

Gould, Eric and Daniele Paserman (2003), "Waiting for mr. right: Rising inequality and declining marriage rates." *Journal of Urban Economics.*

Haider, Steven J., Alison Jacknowitz, and Robert F. Schoeni (2003), "Food stamps and the elderly: Why is participation so low?" *Journal of Human Resources.*

Hampson, Neil B., Christine C. Kramer, Richard G. Dunford, and Diane M. Norkool (1994), "Carbon monoxide poisoning from indoor burning of charcoal briquets." *The Journal of the American Medical Association.*

Hanratty, M. J. (2006), "Has the food stamp program become more accessible? impacts of recent changes in reporting requirements and asset eligibility limits." *Journal of Policy Analysis and Management.*

Hansen, Christian B. (2007a), "Asymptotic properties of a robust variance matrix estimator for panel data when t is large." *Journal of Econometrics.*

Hansen, Christian B. (2007b), "Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects." *Journal of Econometrics.*

Horrace, William and Ronald Oaxaca (2006), "Results on the bias and inconsistency of ordinary least squares for the linear probability model." *Economics Letters*, 90.

Hoynes, Hilary W. and Diane Whitmore Schanzenbach (2009), "Consumption responses to

in-kind transfers: Evidence from the introduction of the food stamp program." *American Economic Journal: Applied Economics.*

Hueglin, Christian, Christian Gaegauf, and Heinz Burscher (1997), "Characterization of wood combustion particles: Morphology, mobility and photoelectric activity." *Environmental Science and Technology.*

Imbens, Guido W. and Michael Kolesar (2012), "Robust standard errors in small samples: some practical advice." *NBER Working Paper.*

Jayachandran, Seema (2009), "Air quality and early-life mortality: Evidence from indonesia's wildfires." *Journal of Human Resources.*

Jeremy Greenwood, Ananth Seshadri and Mehmet Yorukogu (2005), "Engines of liberation." *Review of Economic Studies.*

Johnson, William H. and Jonathan Skinner (1986), "Labor supply and marital separation." *American Economic Review.*

Lam, David (1988), "Marriage markets and assortative mating with household public goods: Theoretical results and empirical implications." *Journal of Human Resources.*

Macaluso, Theodore (2000), "The extent of trafficking in the food stamp program: An update." *USDA Policy Paper.*

Mishra, Vinod, Xiaolei Dai, Kirk R. Smith, and Lasten Mika (2004), "Maternal exposure to biomass smoke and reduced birth weight in zimbabwe." *Population and Health Series.*

Moffitt, Robert (1983), "An economic model of welfare stigma." *American Economic Review.*

National Bureau of Statistics of China (2013a), *China Atmosphere Yearbook*, 2003-2012 edition.

National Bureau of Statistics of China (2013b), *China Health Statistics Yearbook*, 2003-2012 edition.

National Bureau of Statistics of China (2013c), *China Statistical Yearbook*, 1996-2012 edition.

National Bureau of Statistics of China (2013d), *China Urban Construction Statistics Yearbook*, 1996-2012 edition.

Ohls, James and Harold Beebout (1993), *The Food Stamp Program: Design Tradeoffs, Policy and Impacts.* Urban Institute Press, Washington.

Peters, Elizabeth H. (1986), "Marriage and divorce: Informational constraints and private contracting." *American Economic Review.*

Peters, Elizabeth H. (1992), "Marriage and divorce: Reply." *American Economic Review.*

Pope, C. Arden, Joel Schwartz, and Micheal Ransom (1992), "Daily mortality and pm10 pollution in utah valley." *Archives of Environmental Health.*

Rasul, Imran (2006), "The impact of divorce laws on marriage." *mimeo.*

Ritz, Beate and Yu Fei (1999), "The effect of ambient carbon monoxide on low birth weight

among children born in southern california between 1989 and 1993." *Environmental Health Perspectives.*

Siddiqui, Amna R., Ellen B. Gold, Xiaowei Yang, Kiyoung Lee, Kenneth H. Brown, and Zulfiqar A. Bhutta5 (2008), "Prenatal exposure to wood fuel smoke and low birth weight." *Environmental Health Perspectives.*

Stevenson, Betsey (2007), "The impact of divorce laws on investment in marriage-specific capital." *Journal of Labor Economics.*

Stevenson, Betsey and Justin Wolfers (2006), "Bargaining in the shadow of the law: Divorce laws and family distress." *Quarterly Journal of Economics.*

Stevenson, Betsey and Justin Wolfers (2007), "Marriage and divorce: Changes and their driving forces." *NBER Working Paper No. 12944.*

UCLA Center for Health Policy Research (2010), *California Health Interview Survey*, CHIS 2001-2009 adult source file edition. Los Angeles, CA.

U.S. Department of Agriculture (2010), "Food stamp program, year-end participation and bonus coupons issues." *Technical report, Food and Nutrition Service.*

Wang, Qingbin and Qin Zhou (2010), "China's divorce and remarriage rates: Trends and regional disparities." *Journal of Divorce and Remarriage.*

Wang, Xiaobin, Hui Ding, Louise Ryan, and Xiping Xu (1997), "Association between air pollution and low birth weight: A community-based study." *Environmental Health Perspectives.*

Whitmore, Diane (2002), "What are food stamps worth." *Princeton University Working Paper*, 468.

Wolfers, Justin (2006), "Did unilateral divorce raise divorce rates? a reconciliation and new results." *American Economic Review*.

Wright, Richard, Erdal Tekin, Volkan Topalli, Chandler McClellan, Timothy Dickinson, and Richard Rosenfeld (2014), "Less cash, less crime: Evidence from the electronic benefit transfer program." *NBER Working Paper No. 19996*.