2015

Three Essays in Health Economics

Mengcen Qian
Lehigh University

Three Essays in Health Economics

by

Mengcen Qian

Presented to the Graduate and Research Committee
of Lehigh University
in Candidacy for the Degree of
Doctor of Philosophy
in
Business and Economics

Lehigh University
May 2015

Approved and recommended for acceptance as dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Date

_____
Dissertation Advisor

_____
Accepted Date

Committee Members:

_____
Shin-Yi Chou

_____
Muzhe Yang

_____
Mary E. Deily

_____
Suhui Li

Acknowledgements

This dissertation could not have been completed without the encouragement and support of many people. I thank all my committee members for their contributions. I wish to express my deepest gratitude to my advisor and committee chair, Professor Shin-Yi Chou, for her invaluable advice and patient guidance. Her passion and dedication have driven me throughout this entire process. I am thankful to Professors Mary E. Deily, Muzhe Yang, and Suhui Li for their suggestions and comments. In addition, my gratitude is also expressed to Lea R. Gimenez-Duarte and Professor Ernest Lai for their greatest help.

I am thankful to staffs and many nice friends I have made at Lehigh. I would like to thank Rita DiFiore-Czipczer and Josie Sun for editing my paper. I am also thankful to our department coordinator, Rene Hollinger, for her kind assistance. I am especially thankful to my fellow graduates, Lizhong Peng, Xiaohui Guo, Yang Yu, and Krishna Regmi, for accompanying me through years of study in the doctoral program of Economics. I am also grateful to Uncle Bing and Aunt Juan for their generous support for my study life abroad.

It is to my family I owe my greatest appreciation. I am deeply indebted to my parents. Their unconditional love and incredible support give me the strength to overcome difficulties during my pursuit for the degree. I am also grateful to my boyfriend, Xin Zhang, for his endless patience and trust.

Table of Contents

List of Tables

List of Figures

Abstract

In Chapter 1, we use low birth weight (LBW) and intrauterine growth restriction (IUGR) as proxies for a compromised intrauterine environment experienced by one generation, and examine its association with the LBW (or IUGR) status of the next generation. We create two three-generational samples using Taiwan birth certificates from 1978-2006 to study both maternal and paternal transmissions. The results show that the intergenerational transmission only occurs matrilineally and it is stronger among female offspring. We find weak evidence that females, but not males, born to areas with lower unemployment rate, higher average income, and higher parental education can be buffered from these effects.

Chapter 2 uses the most recent vaccine scare in the U.S., the Measles-Mumps-Rubella (MMR)-autism controversy, to investigate how well-educated people respond to information differently when information is mixed. The controversy was first provoked by a paper linking autism to the childhood vaccine of MMR and was retracted years later due to scientific misconduct. We combine state-level information exposures with individual vaccination records from the National Immunization Survey, 1998-2011. Results show that the persistent increase in MMR non-uptake rate is driven by biased beliefs among well-educated mothers, which in turn leads to strong responses to only new information that confirms their beliefs. We find evidence that online search has a more influential impact on the high education group than mainstream media.

Chapter 3 investigates the association between published hospital report card information and hospital relative attractiveness to commercial HMO insurers for CABG surgery in Pennsylvania during report card episodes 2006-2010. Relative bargaining position between hospitals and insurers are measured using aggregated changes in individual willingness-to-pay for a particular

plan if a hospital is included in its network. Plan's hospital networks are implied using PHC4 inpatient discharge data. Our results suggest that high charge hospitals in the most recent report card episode are 53.4% less attractive to insurers and low charge hospitals are 76.1% more attractive to insurers, given the plan's network. Based on our calculation, low charge hospitals in the most recent report card episode are 20.5% more likely to have a new HMO contract.

1. The Intergenerational Transmission of Low Birth Weight: A Large Multigenerational Cohort Study in Taiwan

1.1. Introduction

Prenatal environment has increasingly been recognized as having an important effect on adult health and diseases. Although the link between fetal conditions and future diseases has been studied since the 1940s, it was not until the 1990s that the fetal origins hypothesis was proposed by Barker (1990; 1995). That gave a greater impetus to subsequent research on this temporal linkage. The fetal origins hypothesis states that a fetus faced with a compromised intrauterine environment not only would slow down its growth to reduce nutritional requirements but also might make developmental adaptations by modifying its structure and physiology in a durable fashion, leading to a higher risk of developing chronic diseases in later life. The word "programming" thus is used to describe the linkage between fetal life and long-term consequences (Lucas 1991). Over the past two decades, such an association has been strongly supported by hundreds of human and animal studies.[1] For example, epidemiologic studies in populations worldwide have found that poor fetal growth resulting in low birth weight increases the risk of developing diseases in adulthood, including cardiovascular disease, type 2 diabetes, glucose intolerance, and hypertension. Economic studies further interpret the hypothesis as a major explanation for the temporal relationship between early environment and non-health capital.[2]

More importantly, the evidence--mostly from animal studies--suggests that such impacts of developmental adaptation to the environment occurred during fetal life may not be limited to a single generation but may be transmitted to subsequent generations not exposed to adverse

[1] See Hales and Barker (2001) for a review.
[2] See Almond and Currie (2011) for a review.

environment via non-genomic mechanisms (Drake and Walker 2004; Jablonka and Lamb 2005; Gluckman, Hanson, and Beedle 2007).[3] There are two possible processes underpinning such non-genomic mechanisms. First, adverse *in utero* experiences lead to permanent alterations in physiology, resulting in an adverse intrauterine environment (these includes higher maternal blood pressure, higher maternal insulin, or elevated plasma glucocorticoids during pregnancy) for the fetus, inducing programming effects in the next generation. Second, adverse *in utero* experiences also may influence expression of genes without changing the nucleotide sequences of DNA, through epigenetic modification, thus inducing permanent changes in the phenotype. It is possible to pass an epigenetic trait through both fathers and mothers to the next generation, leading to an intergenerational transmission of fetal programming effects. The latter process highlights the importance of examining the transmission of fetal programming effects, not only matrilineally but also patrilineally. In this paper, we seek to contribute to the scant number of human studies on a general population and to provide evidence on the transmission of programming effects across generations through maternal and paternal lines.

The intergenerational fetal programming effect has important implications. It explains how adverse environmental influences affecting one generation affect the well-being of subsequent generations, a potential mechanism for persistent racial health disparities in the US (Kuzawa and Sweet 2009), or for the so-called "intergenerational cycle of growth failure" in developing countries (Ramakrishnan et al. 1999). Socioeconomic or nutrition interventions that prevent or reverse this transmission could generate positive rewards for future generations. However, it may take the combined efforts of several generations to wash out the impact of an abrupt shock to an

[3] For example, intergenerational inheritance is found in the stress response of both animals and humans, suggesting that there is transgenerational memory of fetal experience that can extend across multiple generations (Matthews and Philips 2010).

ancestor. The effectiveness of such interventions may take longer to manifest, which needs to be taken into account during evaluations.

To examine the intergenerational transmission of the fetal programming effect, we follow most of the literature and use small birth size as a marker of poor fetal nutrition, triggering fetal developmental adjustments that not only slow the growth rate but also influence the future risks of developing chronic diseases. We define small birth size with a low birth weight (LBW, birth weight < 2,500 grams) indicator, which is commonly used in the literature. However, low birth weight is a crude measure for fetal growth, because it can result from prematurity (gestation < 37 weeks), or intrauterine growth restriction (IUGR, also referred to as "small-for-gestational age"), or a combination of the two. Therefore, we further use IUGR as a phenotype of fetal growth. There is no commonly accepted standard definition for IUGR, but the followings are often used: birth weight below the 5th percentile for gestational age; birth weight less than 2,500 grams and gestational age greater than or equal to 37 weeks; and birth weight less than two standard deviations below the mean value of gestational age (Kramer 1987).

Using the annual birth certificates from 1978 to 2006 in Taiwan, we construct three-generational samples. The third generation (G3) includes births that occurred from 1999 to 2006. We then merge data on mothers or fathers for those births to birth certificate data for 1978 to 1985--thus obtaining information on the second generation (G2)--along with demographic information for the grandmother, or the first generation (G1). Identification of G1 is important, because it allows us to control grandmother fixed effects and to net out time-invariant confounding factors, such as shared genes. We create one indicator for LBW and three indicators for IUGR measures for the second and third generations. To examine the intergenerational fetal

programming effects, we study the intergenerational relationships between G2 and G3 for these four markers of fetal growth.

Empirically studying these intergenerational relationships in humans is challenging. The mechanisms underpinning these intergenerational relationships may include not only intergenerational fetal programming effects but also the intergenerational transmission of poverty and shared genes. These intergenerational relationships could be further confounded by the gender of each generation, assortative mating among G2, parenting behaviors (of G1 and G2) after birth (i.e. compensating or reinforcing behaviors), and myriad possibilities of sample selection. Previous epidemiological studies that use natural experiments, such as the Dutch famine during World War II (Lumey 1992), the Chinese famine of 1959-1961 (Fung and Ha 2009), and Ramadan fasting in Tunisia (Alwasel et al. 2013), have provided unique settings for examining the effect of adverse maternal *in utero* environment on offspring growth while netting out potential confounders. However, evidence of such temporal linkages from those studies is mixed, possibly due to small sample sizes and a focus on different cohorts. Small sample sizes make it difficult to obtain precise estimates, and the results obtained from cohorts that experience extreme or specific conditions are difficult to generalize to the entire population.

In this paper, we use within-maternal-sibling-pair or within-paternal-sibling-pair comparisons to estimate the intergenerational correlation in phenotypes of fetal programming from the maternal as well as the paternal side. Our identification strategy is similar to Currie and Moretti (2007) and Royer (2009). The within-G2-sibling comparison allows us to control the genetic predisposition to be small. We include extensive characteristics of mothers and fathers in the regressions in order to control for confounding factors, such as persistent environment, assortative mating of G2, and parental behaviors after birth. We also examine the possible biases

due to sample selection and postnatal investments. We find that sample selection on both G2 and G3 is indeed correlated with their LBW (or IUGR) status. We estimate models accounting for the probability of being observed for the second generation, and find very similar results. We also find little evidence for differential parenting behaviors, suggesting that our results are not due to postnatal investments.

Our contribution to the literature is to address three questions regarding intergenerational fetal programming in humans. First, is LBW (or IUGR) status correlated across generations through maternal and paternal lines? The existing literature mostly focuses on maternal transmission because of missing information on the paternal side. However, as previously mentioned, paternal transmission is possible through epigenetic modification. Moreover, with paternal information we can further control for the confounding effect due to assortative mating. Second, is there a gender-specific effect of such intergenerational transmission? Despite increasing recognition of differential susceptibility to certain outcomes between females and males, few empirical studies focus on a gender-specific pattern in the intergenerational fetal programming effect. Third, can the cycle of intergenerational transmission be modified through interventions that improve socioeconomic status? The second generation in our sample (G2) obtained more years of schooling because of the introduction in 1968 of nine years of compulsory schooling in Taiwan. This generates an arguably exogenous change in socioeconomic status across the generations.

Consistent with the results of epidemiology studies, we estimate a stronger maternal intergenerational transmission on LBW (or IUGR). Other observables do not explain the observed correlations. After controlling for family shared background, the impacts on the paternal side diminish; in contrast, the impacts on the maternal side drop by half, suggesting that shared genetics account for around 50% of the observed maternal correlations. Females are more

affected by this maternal inheritance. Moreover, we find only weak evidence that a child born to a high SES group is less affected by maternal transmission; and such a buffering effect, if there is any, only occurs for females. These findings suggest that maternal health is very important: improving it will provide a healthier intrauterine environment and generate positive spillovers for future generations. Furthermore, socioeconomic interventions may not yield the desired effects within a short period of time. The intergenerational memory of fetal experience may take the efforts of several generations to wash out.

The rest of the paper proceeds as follows. Section 1.2 reviews the literature. Section 1.3 describes the data and constructed samples. Section 1.4 examines whether both maternal and paternal intergenerational correlations in LBW (or IUGR) exist. Section 1.5 investigates the differential inheritance pattern by gender. Section 1.6 examines the potential buffering effects of better socioeconomic status of G2, and section 1.7 concludes.

1.2. Background and Literature Review

1.2.1. Mechanisms for Intergenerational Fetal Programming

Although the biological and molecular mechanisms for intergenerational fetal programming are complicated and not completely understood, two possible pathways have been suggested by existing animal studies. First is the modification of the structure and function of organs and systems involved with metabolism and physiology; second is the modification of the epigenome. We discuss these pathways in more detail below and summarize the discussion in Figure 1.1.

Because the intrauterine environment a fetus experiences is part of the mother's phenotype, the mother's intrauterine environment can influence the intrauterine environment she creates for her offspring. Studies show that those born with reduced size are at increasing risk of developing

hypertension,[4] fetal glucocorticoid overexposure,[5] heightened stress reactivity,[6] and insulin resistance during pregnancy.[7] All four factors, in turn, are strong predictors for LBW or small for gestational age (SGA) of the offspring.

Growing evidence further shows that epigenetic modification may turn out to be the major mechanism for programming that has long-term impacts. It reflects the interacted effects of environment with epigenomes, which can be conceived of as a series of switches that turn on (or off) the expression of various parts of the genome. Fetal life in fact may be the critical stage setting these switches (Petronis 2010; Weaver et al. 2004). Thus, the process gives rise to various phenotypes, even for organisms with the same genetic code.[8] There is some evidence that certain environmentally induced epigenetic markings present in the parent cell can even be maintained during gametogenesis and embryogenesis (Roemer et al. 1997; Morgan et al. 1999), leading to a transgenerational fetal programming effect.

---

[4] Females small for gestational age are at increased risk of developing hypertension during pregnancy (Klebanoff et al. 1999), which in turn predicts low birth weight offspring (Brown et al. 2001; Buchbinder et al. 2002).

[5] Adults born with lower birth weight have elevated plasma glucocorticoids (Phillips et al. 1998; Levitt et al. 2000; Rynolds et al. 2001), which may cause fetal overexposure to maternal glucocorticoids during later pregnancy. Moreover, prenatal glucocorticoid exposure lowers birth weight (McTernan et al. 2001). Biologically, fetal overexposure to glucocorticoids reduces the gene expression of 11β-hydroxysteroid dehydrogenase type 2 (11β HSD2), an enzyme in the placenta that converts active glucocorticoids to inactive products, protecting the fetus from maternal glucocorticoids. Reduced placental 11β HSD2 are found in human pregnancies with intrauterine growth retardation.

[6] Hightened stress reactivity not only restricts fetal growth but also has a direct impact on the develop ment of fetal hypothalamic-pituitary-adrenal (HPA) axis (Worthman and Kuzara 2005). The differences in behavioral and neruoendocrine response to stress can be transmitted across generations (Meaney 2001).

[7] Nutrients delivered across the placenta stimulate fetal production of insulin, a key determinant of fetal growth rate (Lang et al. 2003). Low birth weight is associated with the development of insulin resistance and higher maternal insulin during pregnancy, which further reduces birth size of the offspring.

[8] The epigenetic mechanism modifies gene expressions without changing the nucleotide sequences of DNA. The process gives rise to various phenotypes, typically through DNA methylation or histone modification. Methylation impedes gene expression of that part of DNA to which it is attached. Histone protein can be modified to alter the tightness of DNA packing, thus allowing (or blocking) enzymes and transcription factors to access that stretch of DNA. For example, women with the same genetic code can have different stress reactivity levels based on methylation status of their DNA (Weaver et al. 2004).

Taken together, these two mechanisms seem to suggest a stronger maternal inheritance, because the effect of an adverse intrauterine environment only can be passed on to subsequent generations matrilineally. Nevertheless, the potential transmission through epigenetic modification highlights the importance of examining paternal inheritance as well,[9] although it is largely ignored in the literature.

1.2.2. Intergenerational Correlations in LBW and IUGR

Although birth size has been widely accepted as a major marker for fetal programming (Simon et al. 2006), reduced birth weight may not lie in the causal pathway for disease in adulthood: some gestational exposures are linked to adult disease without any influence on birth size (Morley et al. 2002). To explore more meaningful measures for slow fetal growth *in utero*, we focus on the lower tail of the birth weight distribution: LBW and IUGR. There are no uniform diagnostic criteria for IUGR, so we look at some of the commonly used ones, including SGA at the 5th percentile, less than two standard deviations below mean of gestation, and low birth weight at term.[10] To study the intergenerational fetal programming effects, we examine the intergenerational relationships of these four markers of fetal growth.

However, intergenerational transmission of LBW (or IUGR) may reflect the effects of shared genes and a persistently poor environment. To net out these potential underlying pathways, our fixed-effect models account for the impact of shared family background and genes in the intergenerational correlations. When Currie and Moretti (2007) use a maternal sibling fixed-effect model, they find that LBW women are 50% more likely to deliver LBW infants. The intergenerational transmission of LBW is also stronger for mothers in high poverty zip codes.

[9] For example, the IGF2 gene promotes growth during gestation. Only the allele for IGF2 inherited from the father is expressed (imprinted genes). Imprinted genes may be more susceptible to methylation.
[10] Low birth weight at term is referred as full-term LBW in this paper. It indicates that the birth is at term (between 37 and 42 weeks of gestation) but still weighs less than 2,500 grams.

However, it is a little surprising that the inclusion of sibling fixed effects does not change the coefficients in a noticeable way. Royer (2009) in contrast uses a maternal twin fixed-effect model and finds that a 100-gram increase in maternal birth weight leads to a 7-gram rise in child's weight, a trivial effect. Further, Royer's fixed-effect estimate is roughly 60 percent smaller than the cross-sectional coefficient.[11]

Although twin settings are appealing, in that a twin serves as a near-ideal counterfactual to the other, we do not use them for a couple of reasons. First, we choose to use LBW and IUGR as markers for fetal programming, and the variation of these markers within twin pairs is much smaller than differences in birth weight. Moreover, large differences in intrapair birth weight may indicate a pathologic process that will lead to adverse neonatal outcomes (Hollier, McIntire, and Leveno 1999), resulting in a greater degree of sample selection among twin pairs: G2 twin pairs that make it to the sample will have a lower degree of birth weight discordance. Second, birth weight differences in twins are due mainly to unequal nutrition supply *in utero*, which may be caused by insufficient blood flowing to the placenta in dichorionic twins, or vascular complications of the shared placenta in monochorionic twins. However, epigenetic traits in twins are almost identical during the early years of life (Fraga et al. 2005). Together, these results suggest that if the birth weight difference in twins is through the programming effect, then the most likely underlying mechanism is the physiological change due to undernourishment. Using twin fixed effects thus will preclude the possibility of intergenerational fetal programming from the paternal side through the modification of epigenetic traits. Third, to study the transmission effects from the maternal and paternal sides separately, we can only use same sex twin pairs. This further restricts our sample size, and the variations of LBW (or IUGR) within the twin pairs.

[11] The numbers are taken from Table 3 of Royer (2009) where child's birth weight is the dependent variable. The percentage reduction of the coefficient is (177.87-70.42)/177.87.

Our results may not be generalizable given such a selected sample. Thus, in this paper, we use within-maternal-sibling-pair or within-paternal-sibling-pair comparisons to study the intergenerational transmission effects. We use singleton births for both G2 and G3.

In addition to making a contribution to the literature by studying the transmission effects from mothers and fathers separately, we examine the G3-gender-sepcific effects. Emerging evidence shows that a sex difference in offspring outcomes results from developmental programming. However, the findings on this subject are mixed. Both female-specific (Roseboom et al. 2001; Clifton 2005; Stark et al. 2009) and male-specific (Zaren et al. 2000; Goldenberg et al. 2006; Mingrone et al. 2008) outcomes occur in response to different types of environmental stressors. These results suggest that development in males and females are separate processes from the time of conception (Aiken and Ozanne 2013).

1.3.Data and Sample

We create a maternal and paternal sample using confidential annual birth certificate data for the period 1978-2006 in Taiwan. These data are compiled from forms completed at births and are assembled by the Ministry of Interior. For the entire Taiwan population born during the period, these forms have information on maternal and paternal characteristics (e.g. years of schooling, birth county/town, and birth date), newborns characteristics (e.g. gender, birth order, birth county/town, and birth date), and infant birth outcomes (e.g. birth weight and gestation). Health information at birth is available only in the birth certificate of the particular individual, so we have to match birth certificates over years to obtain birth weight for two generations. We also have personal identification numbers for the infant and both parents, which we use for matching. We focus on singleton births with gestation between 31 and 45 weeks (see footnote 13 for an explanation) and birth weight between 400 and 6,500 grams.

We obtain information for three consecutive matrilineal and patrilineal generations by linking the birth certificates of two generations. For each record in our sample, we have observables for the child (the third generation, G3), the mother or father (the second generation, G2), and the grandmother (the first generation, G1).[12] Our matching procedure is: 1) we set a singleton birth in Taiwan between 1999 and 2006 as the potential G3 in our sample; 2) the birth certificates for singleton births between 1978 and 1985, our G2, are available for us to link; and 3) we merge the birth certificate of the mother (G2) to that of the child (G3) according to mothers birth date and personal identification number to construct a maternal sample. We repeat the procedure to create a paternal sample based on father's birth year and personal identification number. Our final samples provide us with birth weight, gestation age, and other characteristics at birth for G3 and G2, as well as characteristics when giving birth for G2 and G1.

We define LBW for both generations as a dummy equal to one if birth weight is less than 2,500 grams. In order to obtain alternative markers for fetal programming experienced by mother (or father), we estimate birth weight thresholds for SGA at the 5th percentile, denoted as SGA (5th percentile), and at less than two standard deviations below the mean of gestation, denoted as 2SD < mean. To do so, we use the entire singleton population born during 1999-2006 for G3 and during 1978-1986 for G2, respectively.[13] Estimated birth weight thresholds for each gestational age are reported in Appendix Table A1. These thresholds, estimated separately for the two generations, incorporate the impacts of technology changes and medical advances in Taiwan

[12] In the maternal sample, we observe the child's mother and maternal grandmother. In the paternal sample, we observe the child's father and paternal grandmother.

[13] One drawback of using birth certificate to estimate thresholds for IUGR is that we are not able to get credible estimates for very preterm births because there are too few observations as a result of the high rate of fetal mortality. Therefore, we have to limit the lower bound of gestation to 31 weeks in our samples.

over the years.[14] We define SGA (5th percentile) and 2SD < mean dummies equal to one for the child and the mother (or father) if their birth weights are less than the corresponding thresholds. We also define indicators for full-term LBW, denoted as FT LBW, for both generations if their birth weights are less than 2,500 grams at term.[15] Together with LBW, the four indicators are complementary to each other. LBW is determined based on an absolute standard, which is affected by the impact of being preterm. In contrast, SGA (5th percentile) and 2SD < mean are determined based on relative standards for a given gestation. The latter suggests an even more extreme case in birth weight, which is close to the SGA at the 3rd percentile in our samples. However, these two IUGR indicators are potentially subject to measurement errors in gestation. In early years, term infants more often were wrongly recorded as preterm based on the mother's memory of her last menstrual period. In this sense, full-term LBW serves as a better marker among the four, however it is less representative for births before 37 weeks of gestation.

Our third research question, whether the intergenerational transmission could be modified through an improvement in socioeconomic conditions, requires us to measure socioeconomic status at child's birth at an aggregate level to avoid endogeneity. We consider: 1) average income at town-level; 2) unemployment rate at county-level; and 3) average parental education at county-level. We obtain town-level average income at child's birth from Township Income Tax data 1999-2006 provided by the Financial Data Center, Ministry of Finance in Taiwan.[16] We collect county-level unemployment rates by year (1999-2006) from Directorate-General of

[14] It is inappropriate to use fetus growth charts to define indicators for IUGR directly, because most charts are for developed countries, making them inapplicable to the Taiwan population. For example, in some countries the cutoff of 2,500 grams for LBW is about the threshold for SGA at the 10th percentile for birth at 37 weeks of gestation. In contrast, it corresponds to a cutoff for SGA at the 5th percentile for that gestation in Taiwan. It suggests that birth weight exhibits a country-specific pattern.

[15] Births with gestation between 37 and 42 weeks are referred as "at term."

[16] Town-level average income at mother's (or father's) birth is not available. We successfully merged average income during 1999-2006 for 360 towns. Observations from two counties, Jinmen and Lianjiang are excluded. In the year 2004, Middle area and West area are combined as Mid-west area in Tainan city.

Budget under Executive Yuan. We use birth certificates of the entire G3 singleton population to estimate the percentage of at least one parent with years of schooling higher than 9 (or 12) at county-level by year. We also use birth certificates of the entire G2 singleton population to estimate the percentage of G1 with at least 9 years of schooling at county-level at G2's births. Finally, we obtain improvement in education experienced by the child's family, defined as the difference between the above two estimated percentages for at least 9 years of schooling.

There are a total of 280,030 observations in the final maternal sample and 125,078 in the final paternal sample.[17] The paternal sample is less than a half of the maternal sample for several reasons. There is more missing information for the father on the child's birth certificate, which prevents us from tracking father's own birth record. Moreover, our matching procedure requires the second generation to give birth between 1999 and 2006--that is, before the age of 28--in order to be observed in the samples. However, fathers are generally older than mothers at child's birth, resulting in a relatively smaller paternal sample. According to the Demographics Fact Book, Republic of China issued by the Ministry of Interior, the average maternal age at first child's birth is 26.7-28.1 and the average paternal age for having the first child is 30.3-32.9 from 1999-2006. This suggests that our maternal sample is representative, but the paternal sample is relatively young for fathers in Taiwan experiencing births during the G3 period.

We study intergenerational fetal programming by examining the intergenerational correlations in LBW, SGA (5th percentile),[18] 2SD < mean, and FT LBW. Table 1.1 presents the sample means of these four markers of fetal programming for both G2 and G3 in the maternal and

[17] Conditional on birth at term for the child and the mother (or father), the sample size is 255,100 in the maternal sample and 113,369 in the paternal sample.

[18] Another commonly used measure for IUGR is SGA at 10th percentile. In both the maternal and paternal samples, thresholds for SGA at 5th percentile for gestation around 37 weeks are close to 2,500 grams, the cutoff for LBW. Therefore, the results using indicators for SGA at the 5th percentile are more comparable to those using indicators for LBW.

paternal samples. We also provide statistics for a subset sample, conditional on at least one sister (or brother) of the mother (or father) being observed in the samples (denoted as sibling sample). The statistics are similar between the whole and sibling samples, suggesting that our source of variation comes from a subsample that is not selective. We note that the fraction of LBW increased drastically from G2 to G3. There is also a slight increase in the fraction of IUGR as measured by all three criteria, but not as much. Two policy changes are responsible for these trends. First, the birth reporting requirement becomes more stringent after 1994. Before that year, it was common to not report a birth if the newborn was dead. Second, the National Health Insurance (NHI) program implemented in 1995 provides the entire Taiwan population with access to health care at a very low cost. Better medical care allows more preterm births and a weak fetus to survive. In our samples, both policy changes affect the entire third generation, but not the second generation, which explains the observed differences. In section 1.4.4, we account for the potential bias that the probability of observation in samples for the second generation may be correlated with birth weight.

Table 1.2 presents sample statistics of other control variables in the regressions. We note that there is a substantial educational improvement from grandmothers to mothers (or fathers). In 1968, the level of compulsory schooling increased from 6 to 9 years. The mothers and fathers in our samples were affected by that policy. In contrast, only about 0.8% of the G1 in our samples were affected by it. Roughly 10% of grandmothers have more than 9 years of schooling. However, that number skyrockets to 80% for G2. This drastic change is due, more or less, to the 1968 compulsory education law, which provides us with an arguably exogenous change in socioeconomic status across generations. We discuss this in more detail in section 1.6.

1.4. Intergenerational Correlation in LBW

1.4.1. Estimation Strategy

To estimate the intergenerational correlation, we assume that a child's marker of fetal programming, such as LBW, is an additively separable linear function of the mother's (or father's) marker and a matrilineal (or patrilineal) family fixed effect.[19] For each child i of mother (or father) j of grandmother k, we consider a grandmother fixed-effect model

$$LBW_{ijk}^{G3} = \alpha_k + \beta_{FE} LBW_{jk}^{G2} + \gamma\left(X_{ijk}^{G3}, X_{jk}^{G2}, X_k^{G1}\right) + \delta_1 edu_{spouse}^{G2} + \delta_2 age_{spouse}^{G2} + \varepsilon_{ijk}, \quad (1.1)$$

where LBW_{ijk}^{G3} is a dummy equal to one if the child is LBW; LBW_{jk}^{G2} is the key independent variable, an indicator for LBW of the mother (or father); $\left(X_{ijk}^{G3}, X_{jk}^{G2}, X_k^{G1}\right)$ is a vector of observables for all three generations; edu_{spouse}^{G2} and age_{spouse}^{G2} are years of schooling and age of the spouse of the second generation at child's birth, which attempts to capture the mating behavior of the second generation;[20] α_k is the grandmother fixed effects, representing the time-invariant heterogeneity within the family; and ε_{ijk}, is an idiosyncratic error term. Grandmother fixed effects are used to capture the genetic factors that are common to maternal or paternal siblings, but also indicate other shared family background, such as consistent health behaviors and parenting styles. We could also have controlled for grandfather fixed effects in both samples. In essence, both types of fixed effects account for the same source of shared unobservables on

[19] The specification is less flexible in the sense that it does not allow a gene and environment interaction (Royer and Witman, 2013).

[20] We could obtain more information on the spouse of the second generation by merging the maternal sample to the paternal sample. However, such a combined sample suffers from a severe selection issue because only children with both parents below the age of 28 will be observed.

maternal or paternal line.[21] In equation (1.1), β_{FE} is the coefficient of interest. Its impact is identified by variations in LBW among children whose mothers (or fathers) are sisters (or brothers). If the fetal programming effect is inheritable, then we would expect the sign of the coefficient to be positive and significant. We also run regressions replacing the LBW indicator in equation (1.1) with three indicators of IUGR for both G3 and G2. The standard errors are clustered at child's hospital-year level.

One limitation of the model is that we cannot capture family-specific time-varying variables; this may lead to different outcomes among the third generation of the family. Following Currie and Moretti (2007), we therefore add observables, step by step, to examine the impact of other confounding factors on our estimates. First, we estimate an OLS specification without other controls. Next, we add gestation dummies for pre-term, at term, and post-term. Two factors can explain extremes in birth weight: being preterm and the growth rate at a fixed gestation. Our specification using gestation dummies is expected to net out any variations in the child's LBW that come from being pre-term. In an additional step, we add variables that we treat as predetermined before the birth of G3. They include: dummies for G2 and G3's birth year, to account for trends in birth weight over time; dummies for G2's birth order (first, second, or third) and birth place (hospital, or clinics and maternity homes); dummies for G1's years of schooling (7-9 years, 10-12 years, 13-14 years, 15-16 years, or above 17 years); age (21-25, or 26-30) and marital status when giving child birth; and interactions between dummies for G1's county of residence at G2's birth and G2's birth year.[22] Then, we include variables that are not strictly

[21] In both samples, only around 1.1% of the grandmothers gave birth with different spouses. Most of those cases resulted from typos in grandfather's personal identification numbers. Therefore, switching from grandmother fixed effects to grandfather fixed effects in both samples generates similar results.

[22] G1's county of residence at G2's birth is assumed to be the county where G2 was born, indicated in G2's birth certificate. Similarly, mother's county of residence at child's birth is assumed to be the county where the child was born, indicated in child's birth certificate.

exogenous, in the sense that they may be jointly determined with the decision of giving birth: dummies for child's birth order; G2's years of schooling, age, and marital status at child's birth; interactions between dummies for G2's county of residence at child's birth and child's birth year; and town-level average income at child's birth. Finally, we account for assortative mating by adding dummies for G2 spousal years of schooling and age at the child's birth,[23] and then control grandmother fixed effects.

1.4.2. Results

Tables 1.3 and 1.4 present the intergenerational correlations in LBW, SGA (5th percentile), 2SD < mean, and FT LBW using the maternal and paternal sample, respectively. In both tables, the additional controls just described are added in turn from columns (1) to (6). Without controlling any additional variables, a child born to a LBW mother is 5.40 percentage points more likely to be LBW (column (1) of Table 1.3). In contrast, a child born to a LBW father is only 2.25 percentage points more likely to be LBW (column (1) of Table 1.4). The estimates for SGA (5th percentile), 2SD < mean, and FT LBW all yield similar patterns: maternal correlation is stronger than paternal correlation. Adding additional controls does not change the estimates very much in columns (1) to (5) of both tables. This suggests that observables do not explain much about intergenerational transmission. In the maternal sample, the magnitudes of all correlations decline and remain significant when we add grandmother fixed effects in column (6). Comparing those to the estimates in column (5), we find that shared genes explain about a half of the correlation. This suggests that genetics or shared background is an important determinant of intergenerational correlation of LBW (or IUGR), in contrast to the findings of Currie and Moretti (2007). Accounting for grandmother specific effects, a child born to a LBW mother is around 36%

[23] In the maternal sample, there are six dummies for spousal age (21-25, 26-30, 31-35, 36-40, 41-45, and 46-50) in the paternal sample, there are five dummies for spousal age (21-25, 26-30, 31-35, 36-40, and 41-45).

more likely to be LBW (column (6) of Table 1.3).[24] This impact is slightly smaller than the estimated 50% that Currie and Morretti (2007) find. However, we find no similar evidence in the paternal sample. In column (6) of Table 1.4, all four estimates lose significance after we control for paternal grandmother fixed effects. For the models with SGA (5th percentile) and 2SD < mean, the estimated correlations become negligible.

Because results for IUGR indicators may be contaminated by measurement errors in G2 gestational age, we regress four markers of G3 on G2's LBW status. We find similar results: intergenerational correlations are stronger for the maternal sample, and controlling grandmother fixed effects reduces the correlations by about half. These estimates are reported in Tables A2 and A3 in Appendix A.

After accounting for shared genes, the stronger maternal intergenerational correlations in LBW (or IUGR) are consistent with the results of biology studies (Magnus et al. 2001; Collins et al. 2002; 2003; Kuzawa and Sweet 2009). Part of the intrauterine environment that the fetus experienced is an expression of the maternal phenotype. The programming effect experienced by mothers, as reflected in LBW (or IUGR), serves as a signal for the fetus to make developmental adaptations without the presence of other environmental stressors. Along with the additional pathways discussed in section 1.2.2, the consequence of maternal fetal programming appears to be more durable. This suggests that improving maternal health will generate a positive spill-over effect on offspring.

1.4.3. Postnatal Investments by G1

The difference in maternal and paternal transmission presented in Tables 1.3 and 1.4 does not account for possible differential parental investments by gender. Postnatal investment by

[24] The number is obtained using the baseline incidence of LBW for G3 in the maternal sample, 0.0228/0.0626=36.4%.

grandparents over the childhood of mothers and fathers may change their health conditions at the time they give birth to the third generation, and thus obfuscate the biological impact of intergenerational correlation in LBW (or IUGR). For example, if parents invest more heavily in a disadvantaged male child, their behavior may prevent us from observing a correlation in LBW between the child and the father. On the other hand, if parents care less for a disadvantaged female child because of potentially low returns to their investment, their behavior may lead to a stronger correlation in LBW on the maternal side. Under these scenarios, our findings would be driven by grandparents allocating resources differently by gender, based on G2's LBW (or IUGR). The estimated effects would be upward biased in the maternal sample and downward biased in the paternal sample.

We cannot estimate the relationship between grandmother's parenting behavior and G2's LBW (or IUGR) status using birth certificates alone because they lack such information. So following Royer (2009), we instead examine whether intergenerational correlations differ across families whose ability to invest more on weak children varies. For example, large families may have a tighter budget and be less likely to invest more on one particular child. In all the regression models, we add to the most inclusive specification presented in Tables 1.3 and 1.4 an interaction term between G2's LBW (or IUGR) and an indicator equal to one if G2 was born into a large family. We define a large family as the mother (or father) having at least two older siblings. These results are reported in Table A4 in Appendix A. We find that the differential maternal correlations in LBW (or IUGR) are generally negligible in large families relative to small families. Our results suggest that the strong evidence for maternal transmission of LBW (or IUGR) found in Table 1.3 is not driven by differential parenting behavior of the grandmother.

1.4.4. Sample Selection on G2

Conditional on the survival of G3, the mother and father will be observed in our samples if they give birth to a singleton between 1999 and 2006. Therefore, selection on G3 and selection on G2 may bias our estimates. A severe intergenerational fetal programming effect can lead to fetal mortality, thus preventing us from observing G3 from birth certificates. This type of selection will lead us to underestimate the intergenerational effect, because those influenced by the intergenerational fetal programming will not be in the sample. The impact of selection on G2 is ambiguous, though. On one hand, weaker G2 may not be observed in our samples due to mortality, inferior marriage market outcome, or delay in fertility. On the other hand, stronger G2 may not be observed in our samples because of higher educational attainment that delays marriage and fertility (Almond 2006). In the former case, our estimates will be understated; in the latter case, they will be overstated.

To get a sense of the potential bias due to sample selection, we run a regression using the entire G2 sample, with and without having G3 offspring. The dependent variable $InSample$ is equal to one if G3 offspring is observed; that is, the G2-G3 pair is in our analysis samples. We estimate the probability of being observed in our analysis samples using the following equation:[25]

$$InSample_{jk}^{G3} = \alpha_k + \beta_{FE} LBW_{jk}^{G2} + \gamma\left(X_{jk}^{G2}, X_k^{G1}\right) + \varepsilon_{jk}. \quad (1.2)$$

In Table 1.5, we show that the second generation born with reduced birth size--measured by birth weight and indicators for LBW, SGA (5th percentile), 2SD < mean, and FT LBW--is less likely to be observed in both the maternal and paternal samples. Our fixed-effect estimates suggest that

[25] Observed controls include dummies for G2's birth year, birth order, birth place and gestation; dummies for G1's years of schooling, age, marital status, and county of residence at G2's birth; and interactions between dummies for G1's county of residence at G2's birth and G2's birth year.

LBW mothers are 1.26 percentage points less likely to be observed in the maternal sample and LBW fathers are 1.22 percentage points less likely to be observed in the paternal sample. To further gauge how much this sample selection may bias our results, we provide two sets of robustness checks. First, we perform a series of nonparametric tests following Royer (2009). Then, we include the probability of being observed by using different functional forms in estimating equation (1.1).

The basic idea of the "nonparametric" test is that, for observations in groups with different degrees of sample selection, if the intergenerational correlations are identical across the groups then sample selection bias may not be an issue. To carry out this test, we divide G2 into groups based on available observables, such as birth cohort of G2 and G1's years of schooling. We then test whether the effect of LBW on the probability of later observation differs across groups. In other words, we include the interaction terms between G2 LBW and categorical observables (listed in Table A5) in equation (1.2), and then perform a joint F-test on those interaction terms. If the interaction terms are jointly significant, this implies that sample selection based on the observable occurs. For the observables that do lead to sample selection across groups, we further test whether the intergenerational correlations in LBW (or IUGR) are identical across groups by including the same interaction terms into equation (1.1). We perform these two-step tests separately by maternal and paternal sample. The results, presented in Table A5 in Appendix A, suggest that the effects of LBW on the probability of being observed do not vary by most of the observables. Out of eight joint tests, only one is statistically significant at the 1% level for the maternal sample; three are statistically significant at the 5% level for the paternal sample. Table A6 shows the results of including the interaction terms between LBW and the observables that

are jointly significant in the first step into equation (1.1). None of the joint tests are statistically significant, suggesting that sample selections based on observables do not bias our estimates.

The results adjusted for the predicted probability of being observed, denoted as $\hat{p}(x)$ as estimated by equation (1.2), are reported in Tables 1.6 and 1.7 for the maternal and paternal samples, respectively. Column (1) of both tables presents our original fixed-effect estimates as shown in column (6) of Tables 1.3 and 1.4. Columns (2) to (5) report estimates after our four adjustments: the estimates from a weighted regression using inverse probability, $1/\hat{p}(x)$, are reported in column (2); the estimates controlling directly for the probability are presented in column (3); the estimates controlling for a quadratic form of probability are displayed in column (4); and the estimates from models capturing selection by adding the interaction between G2 LBW (or IUGR) and demeaned estimated probability of being observed, LBW x $(\hat{p}(x) - \overline{\hat{p}(x)})$ are reported in column (5). We also report the p-values for t-tests on the equality between the adjusted estimates in column (2) to (5) and those in column (1).

After accounting for the probability of being observed in the sample, our results are largely the same. With large p-values, we are unable to reject the null hypothesis that adjusted and unadjusted estimates are the same for all outcomes. These results suggest that our findings are robust after accounting for sample selection on G2, which is consistent with the results from the earlier "nonparametric" tests.

1.5. Differential Inheritance Patterns by Gender

1.5.1. Estimation Strategy

After finding no significant impact from the paternal sample, we turn our focus to the maternal intergenerational effect of LBW (or IUGR). To estimate the differential inheritance

patterns by gender of G3, for each child i of mother j of grandmother k, we consider the following grandmother fixed-effect model:

$$LBW_{ijk}^{G3} = \alpha_k + \beta_{FE} LBW_{jk}^{G2} + \beta_{male} LBW_{jk}^{G2} \text{ x } male + \gamma\left(X_{ijk}^{G3}, X_{jk}^{G2}, X_k^{G1}\right) + \delta_1 edu_{spouse}^{G2} + \delta_2 age_{spouse}^{G2} + \varepsilon_{ijk}, \quad (1.3)$$

where male is a dummy equal to one for male birth; the other controls are the full set of observables held constant in the regressions as in equation (1.1).[26] β_{male} is the coefficient of interest; it captures the differential impact on male birth from maternal transmission.

1.5.2. Results and Discussions

Table 1.8 presents the estimates of differential maternal transmission by gender on all four outcomes. The estimates from the model with LBW show that difference by gender is small and insignificant. However, maternal inheritance of SGA (5th percentile) and FT LBW is significantly smaller for male than for female offspring. The results from these two outcomes suggest that female infants born to IUGR mothers are 50-70% more likely to be IUGR than those not born to IUGR mothers. In contrast, male infants born to IUGR mothers are 20-50% less likely to be IUGR than those not born to IUGR mothers.[27] This evidence suggests stronger maternal transmission in fetal growth for females, which is consistent with the findings from similar animal studies. Observing rhesus monkey across several generations, for example, Price and his colleagues find that intergenerational correlation in fetal growth has followed a matrilineal pattern and was much more pronounced for female than for male offspring (Price, Hyde, and Coe 1999; Price and Coe 2000).

[26] Main effect of child's gender is included in other controls.

[27] In the maternal sample, the base-line incidences of SGA (5th percentile) for G3 are 0.082 for female birth and 0.046 for male birth; the base-line incidences of 2SD < mean for G3 are 0.029 and 0.016 for female and male births, respectively.

Explanations from evolutionary biology provide a way to understand the observed sex difference in maternal transmission of LBW. To have the best chance of reproductive success for the overall species, it may be more effective and efficient for mothers to invest heavily in the long-run protection of their female fetuses, thus making female offspring more sensitive but also more adaptable to the intrauterine environment (Aiken and Ozanne 2013). Mothers LBW (or IUGR) thus may serve as an integrated signal, reflecting recent intrauterine environments experienced by matrilineal ancestors (Kuzawa 2005). Therefore, the observed stronger maternal transmission for females would enable them to make more stable adaptations, filtering out the noise of potential short-term fluctuations in environmental stress. Even though fetuses of both sexes are affected by any given stress, the process could be experienced differently for females and males.[28]

1.5.3. Parenting Behavior by G2

If mothers perform different parenting behaviors when they find out the gender of the unborn child, for example, taking better care of themselves if they know they are carrying a boy, then there would be an upward bias in the observed differential impact among males. We cannot however identify mothers who know the gender of their child before birth in our dataset. As in section 1.4.3, we instead examine gender differences in intergenerational correlations across families with varying ability to invest more in the unborn child. We define a large family as the child having at least two older siblings, and we expect large families to be less likely to treat one unborn child better than another because of limited resources. We add a triple interaction term to equation (1.3): mother's LBW (or IUGR), male birth indicator for the child, and a dummy equal to one if the child is born to a large family. The results are reported in Table A7 in Appendix A.

[28] For example, hypertension in men was linked to the mother's socioeconomic status, an indicator of their diets; in contrast, hypertension in women was linked to the mothers height, an indicator of her protein metabolism (Eriksson et al. 2010).

Out of four tests, we only find a positive and significant coefficient on this triple interaction term in the model with 2SD < mean. Our results provide only weak evidence that mothers may perform compensating behaviors to male births if boys are small.

1.5.4. Sample Selection on G3 by Gender

Greater environmental adaptability in female fetuses leads to more stable reproductive outcomes for female offspring. Thus, female and male fetuses affected by an intergenerational fetal programming effect may encounter different degrees of fetal mortality risk, potentially leading to stronger maternal transmission in LBW (or IUGR) on female births (See Appendix B for the proof). Similarly, the Triver-Wilard hypothesis predicts that mothers in poor conditions have more daughters, because of female having a greater chance of reproductive success (Trivers and Wilard 1973). Both explanations indicate that the probability of observation for a female child in the sample could be significantly greater than that for a male child if the mother is LBW (or IUGR), and this may be driving the observed gender difference.

To measure the potential bias due to sample selection on child by gender, we estimate the impact of mothers LBW (or IUGR) on the probability of observing a male birth in the maternal sample, using the most inclusive specification of equation (1.1).[29] The results are presented in Table A8 in Appendix A. We find that all the estimates of fetal programming markers (LBW, SGA (5th percentile), 2SD < mean, and FT LBW) are small and insignificant, suggesting that sample selection by gender is not likely to bias our findings.

[29] Gender of G3 will be excluded from the explanatory variables.

1.6. Differential Maternal Transmission by SES

1.6.1. Estimation Strategy

We use a difference-in-difference-in-difference model to test whether the intergenerational transmission can be buffered by socioeconomic interventions and whether such a protective effect also exhibits a gender-specific pattern. For each child i of mother j of grandmother k, we consider the following grandmother fixed-effect model

$$LBW_{ijk}^{G3} = \alpha_k + \beta_1 LBW_{jk}^{G2} + \beta_2 male + \beta_3 highSES + \beta_4 LBW_{jk}^{G2} \text{ x } male +$$

$$\beta_5 LBW_{jk}^{G2} \text{ x } highSES + \beta_6 male \text{ x } highSES + \beta_7 LBW_{jk}^{G2} \text{ x } male \text{ x } highSES +$$

$$\gamma\left(X_{ijk}^{G3}, X_{jk}^{G2}, X_{k}^{G1}\right) + \delta_1 edu_{spouse}^{G2} + \delta_2 age_{spouse}^{G2} + \varepsilon_{ijk}, \quad (1.4)$$

where $highSES$ is a dummy equal to one if the child is born into a high socioeconomic group. This is defined based on: 1) average town-level income at G3's birth above the mean value of 1999-2006; 2) county-level unemployment rate at G3's birth above the mean value of 1999-2006; and 3) county-level percentage of at least one parent of G2 with above 12 years of schooling at G3's birth above the mean value of 1999-2006.[30] Finally, we use the change in educational attainment between G1 and G2 to measure the improvement in SES. As mentioned in the Introduction, we exploit the arguably exogenous change in SES across the generations due to the introduction of compulsory schooling in Taiwan in 1968. We define $highSES$ as a dummy equal to 1 if the difference in county-level percentage of at least 9 years of schooling from G1 to G2 is above the mean value of the sample. In equation (1.4), β_5 is the coefficient of interest. It captures

[30] We divide the high income group based on the mean of town-level average income at child's birth instead of the median because income distribution is skewed. For the rest of the measures for high SES, we use the mean as the threshold for consistency in reporting; the results are unaffected by changing the cutoff to the median. Moreover, the criterion for high education for G2 being higher than that for G1 at the time of giving birth is due to the policy change regarding compulsory schooling in 1968.

the differential impact in the high SES, or most improved, group for females. The sum of β_5 and β_7 is the differential impact in the high SES, or most improved, group for males.

1.6.2. Results and Discussions

Table 1.9 presents the estimates for β_1, β_4, β_5, and β_7 in equation (1.4). Based on the p-values for the t-test (row (f) in each panel), we are unable to reject the null hypothesis that there is no differential impact of G2 LBW (or IUGR) on male children born to the high SES group in all panels across all models. In contrast, we find some evidence that females born to the high SES groups are less affected by the intergenerational correlation in LBW (or IUGR). Out of 16 coefficients (row (c) in each panel), three coefficients in panel A and one coefficient in panel B and panel C are statistically significant at the 5% level (two of these five are significant at the 1% level). In panel A, except for SGA (5th percentile), females born to LBW (or IUGR) mothers in a county with a low unemployment rate are 2.26-2.50 percentage points less likely to be LBW (or IUGR). This difference represents a decrease of around 30% as compared to the base-line incidence of LBW (or IUGR) in females. The evidence from town-level income and parental education is weaker. However, we only find a significant differential impact on females born into towns with high average income in the model with 2SD < mean (in panel B) and those born in counties with high parental education in the model with FT LBW (in panel C).[31] We find no differential impact on males and females born in counties that experienced the most improvement in SES (in panel D). Thus, our results weakly support the findings in the literature: children born into favorable socioeconomic conditions suffer less as a result of poor maternal health (Currie and Moretti 2007; Bhalotra and Rawlings 2013). Moreover, our findings indicate

[31] We also measured the socioeconomic status using G1's education level at mother's birth. We define a child as born to a high SES group if the percentage of G1's years of schooling above 9 at the county-level at the time of the mother's birth is above the mean. Although the variation in SES at birth among mothers who are siblings is rather limited, we do find weak evidence that females, but not males, born to the high SES group suffer less from the maternal transmission of LBW (or IUGR).

that such a buffering effect only occurs for females, which may be attributable to the greater sensitivity of females to the maternal intrauterine environment.

Although the evidence for such a buffering effect is weak, it clearly suggests that creating a less stressful living environment for mothers--especially during a critical stage of life such as pregnancy--will mitigate the intergenerational transmission of maternal poor health for females, possibly through improving the intrauterine environment. It may take the collective effort of several generations to completely wash out the programming effect in a given matrilineal line, but the rewards for females will generate positive spill-over effects to future generations.

1.6.3. Sample Selection on G3 by SES and Gender

If the probability of observation for a male birth is different from that for a female birth for the third generation (G3) by socioeconomic group, this may potentially drive our observed difference in the buffering effect of SES intervention by gender. As in section 1.5.4, we therefore regress the indicator for male birth on mother's LBW (or IUGR) and an interaction term between mother's LBW (or IUGR) and dummies for high SES measured at child's birth. Our results are reported in Table A9 in Appendix A. We do not find that sample selection on G3 differs by gender across SES groups as measured by all four criteria in all outcomes.

1.7. Conclusion

This paper uses two three-generational samples of Taiwan-born singletons to estimate the intergenerational transmission of LBW (or IUGR) from the maternal and paternal side, respectively. The intergenerational fetal programming effect provides the biological mechanism for such correlation. We use LBW (or IUGR) as markers for experiencing an adverse intrauterine environment. We use grandmother fixed effects to examine how an unfavorable *in utero*

experience for parents may pass through to their offspring. This is appealing, because it controls for unobserved heterogeneity across families and incorporates evidence from both genders.

We find that the intergenerational correlation of LBW (or IUGR) only occurs matrilineally. Specifically, children born to LBW mother are 36% more likely to be LBW, after accounting for shared family background. In contrast, there is no significant inheritance on the paternal side. Further, such correlation is stronger in female offspring when we use IUGR measures. We find only weak evidence that intergenerational transmission of LBW (or IUGR) is buffered by high SES. Moreover, such a buffering effect, if it exists, is only found for female offspring. Based on several robustness checks, we conclude that our results are not driven by sample selection or by differential parenting behaviors by gender.

These findings suggest that maternal health is very important because the consequences of exposure to an adverse *in utero* event can extend to multiple future generations through the matrilineal line via a non-genomic mechanism. Socioeconomic improvements have only weak ameliorative effects on this intergenerational transmission. Therefore, it may take the collective effort of several generations to wash out the transgenerational memory of an unfavorable fetal experience. A longer study window may be more appropriate for evaluating the effectiveness of interventions that focus on the wellbeing of the mother.

There are some caveats in our study. First, because of the matching procedure, the paternal sample is smaller and the fathers observed are generally younger than average for the second generation in Taiwan at the time of giving birth. This makes some of our results less precise and less representative in the paternal sample. Second, extremes in birth weight are only broad measures for intrauterine environment. Biological measures, such as insulin sensitivity, blood pressure, and stress response, may better capture the biological mechanism underlying the

intergenerational fetal programming effect on specific health outcomes. This could be a focus for future research.

Figure 1.1 Underlying Mechanisms of Intergenerational Inheritance of Fetal Programming

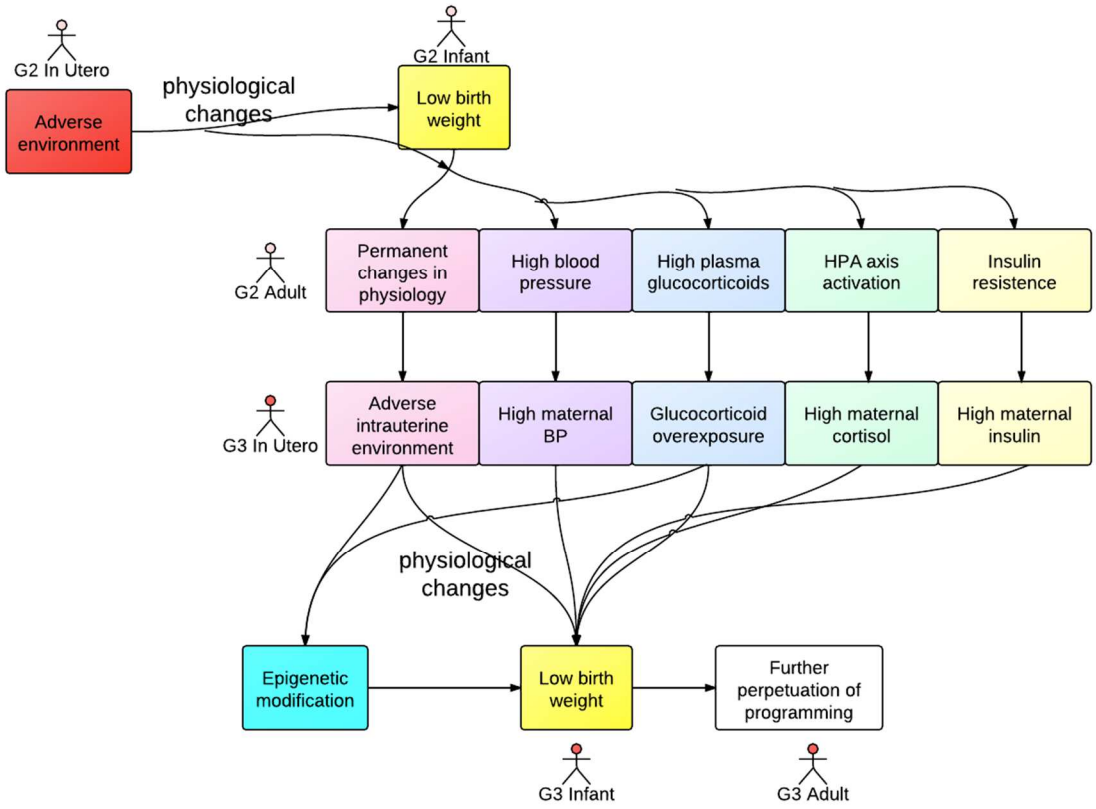


Table 1.1 Sample Means for Markers of Fetal Programming for G2 and G3

	Maternal sample		Paternal sample	
	Whole sample	Sibling sample[1]	Whole sample	Sibling sample[1]
Markers for G3 (1999-2006)				
LBW	0.0626	0.0662	0.0659	0.0715
SGA (5th percentile)	0.0635	0.068	0.0663	0.0697
2SD < mean	0.0225	0.0246	0.0234	0.0257
FT LBW[2]	0.0367	0.0400	0.0386	0.0402
Markers for G2 (1978-1985)				
LBW	0.0363	0.034	0.0258	0.0274
SGA (5th percentile)	0.0588	0.0575	0.0367	0.0386
2SD<Mean	0.0233	0.0222	0.0148	0.0169
FT LBW[3]	0.0275	0.0264	0.0179	0.0196
Sample size	280,030	46,849	125,078	9,181

[1] Maternal sibling sample includes mothers that have at least one sister in the sample and paternal sibling sample includes fathers that have at least one brother in the sample.

[2] For full-term LBW, the sample sizes for the maternal samples are 261,478 for the whole sample and 43,555 for the sibling sample. The sample sizes for the paternal samples are 116,509 for the whole sample and 8,527 for the sibling sample.

[3] For full-term LBW, the sample sizes for the maternal samples are 273,109 for the whole sample and 45,717 for the sibling sample. The sample sizes for the paternal samples are 122,006 for the whole sample and 8,933 for the sibling sample.

Table 1.2 Sample Means of Other Control Variables

	Maternal sample		Paternal sample	
	Whole sample	Sibling sample	Whole sample	Sibling sample
G3 characteristics				
Gestational age				
Preterm (33-36 weeks)	0.0620	0.0661	0.0643	0.0651
Full-term (37-42 weeks)	0.9337	0.9297	0.9315	0.9288
Post-term (43-45 weeks)	0.0005	0.0005	0.0004	0.0008
Birth order				
First born	0.6459	0.6012	0.7003	0.6595
Second born	0.2970	0.3241	0.2595	0.2885
Third born	0.0497	0.0645	0.0357	0.0457
Birth place				
Hospital	0.6024	0.5912	0.5924	0.5752
Clinics or maternity homes	0.3972	0.4084	0.4071	0.4245
G2 (mother or father) characteristics				
Birth order				
First born	0.3190	0.2448	0.3080	0.2744
Second born	0.2974	0.3157	0.2909	0.3446
Third born	0.2192	0.2515	0.2260	0.2486
Birth place				
Hospital	0.6687	0.6115	0.6542	0.6019
Clinics or maternity homes	0.3122	0.3653	0.3261	0.3757
Married	0.9859	0.9846	0.9860	0.9845
Age at child birth				
21-25	0.6205	0.6383	0.5954	0.6456
26-30	0.2083	0.1464	0.3089	0.2290
Years of schooling at child birth				
7-9 years	0.1637	0.2158	0.2006	0.2525
10-12 years	0.7245	0.7132	0.6992	0.6852
13-14 years	0.0514	0.0304	0.0447	0.0298
15-16 years	0.0430	0.0210	0.0381	0.0184
> 16 years	0.0025	0.0007	0.0060	0.0016
G1 (grandmother) characteristics				
Married	0.9857	0.9879	0.9857	0.9863
Age at child birth				
21-25	0.4801	0.5152	0.4852	0.5474
26-30	0.2733	0.2467	0.2777	0.2185
31-35	0.0537	0.034	0.0533	0.0200
36-40	0.015	0.0062	0.0152	0.0038

41-45	0.004	0.0014	0.0040	0.0011
Years of schooling at child birth				
7-9 years	0.1981	0.1893	0.1813	0.1842
10-12 years	0.0953	0.0553	0.0925	0.0542
13-14 years	0.0037	0.0013	0.0041	0.0017
15-16 years	0.0016	0.0004	0.0019	0.0005
> 16 years	0.0005	0.0006	0.0004	0.0001
Sample size	280,030	46,849	125,078	9,181

Table 1.3 Effect of Mother's LBW (or IUGR) on Child's LBW (or IUGR)[1]

	(1)	(2)	(3)	(4)	(5)	(6)
Sample: gestation ≥ 31 weeks (N=280,030)						
Impact of G2 LBW on G3 LBW	0.0540***	0.0427***	0.0428***	0.0424***	0.0423***	0.0228***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)
Impact of G2 SGA (5th pctl.) on G3 SGA (5th pctl.)	0.0565***	0.0568***	0.0571***	0.0568***	0.0566***	0.0225***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)
Impact of G2 2SD<mean on G3 2SD<mean	0.0298***	0.0299***	0.0300***	0.0296***	0.0295***	0.0192***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)
Sample: gestation between 37 and 42 weeks (N=255,100)						
Impact of G2 FT LBW on G3 FT LBW	0.0428***	0.0428***	0.0428***	0.0423***	0.0422***	0.0219***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.008)
G3 gestational age		Y	Y	Y	Y	Y
Pre-determined variables before G3 birth[2]			Y	Y	Y	Y
Jointly-determined variables with pregnancy of G3[3]				Y	Y	Y
G2 spousal characteristics[4]					Y	Y
G1 (grandmother) fixed effects						Y

[1] Standard errors are clustered at child's hospital and year level in parentheses. Each column of each row is a separate regression. All dependent variables are dummies for G3's LBW (or IUGR). Coefficients are reported for indicators of G2's LBW (or IUGR). *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

[2] Pre-determined variables before G3 birth include G3's gender, birth year dummies, G2's birth order dummies (first, second, and third), birth place dummies (hospital and clinics or maternity homes), G1's marital status at G2's birth, and interactions between dummies for G1's county of residence at the time of G2's birth and dummies for G2's birth year.

[3] Jointly-determined variables with pregnancy of G3 include G3's birth order dummies (first, second, and third), G2's five education dummies, two age dummies and marital status at child's birth, town-level average income at child's birth year, and interactions between dummies for G2's county and dummies for G3's birth year.

[4] G2 spousal characteristics include five dummies for spousal education and six age dummies.

Table 1.4 Effect of Father's LBW (or IUGR) on Child's LBW (or IUGR)[1]

	(1)	(2)	(3)	(4)	(5)	(6)
Sample: gestation ≥ 31 weeks (N=125,078)						
Impact of G2 LBW on G3 LBW	0.0225***	0.0204***	0.0200***	0.0196***	0.0194***	-0.0123
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.022)
Impact of G2 SGA (5^{th} pctl.) on G3 SGA (5^{th} pctl.)	0.0337***	0.0337***	0.0331***	0.0326***	0.0323***	0.0093
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.020)
Impact of G2 2SD < mean on G3 2SD < mean	0.0145***	0.0145***	0.0138***	0.0134***	0.0133***	0.0086
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.019)
Sample: gestation between 37 and 42 weeks (N=113,639)						
Impact of G2 FT LBW on G3 FT LBW	0.0269***	0.0269***	0.0262***	0.0256***	0.0255***	0.0294
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.024)
G3 gestational age		Y	Y	Y	Y	Y
Pre-determined variables before G3 Birth[2]			Y	Y	Y	Y
Jointly-determined variables at G3 Birth[3]				Y	Y	Y
G2 spousal characteristics[4]					Y	Y
G1 (grandmother) fixed effects						Y

[1] Standard errors are clustered at child's hospital and year level in parentheses. Each column of each row is a separate regression. All dependent variables are dummies for G3's LBW (or IUGR). Coefficients are reported for indicators of G2's LBW (or IUGR). *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

[2] Pre-determined variables before G3 birth include G3's gender, birth year dummies, G2's birth order dummies (first, second, and third), birth place dummies (hospital and clinics or maternity homes), G1's marital status at G2's birth, and interactions between dummies for G1's county of residence at the time of G2's birth and dummies for G2's birth year.

[3] Jointly-determined variables with pregnancy of G3 include G3's birth order dummies (first, second, and third), G2's five education dummies, two age dummies and marital status at child's birth, town-level average income at child's birth year, and interactions between dummies for G2's county and dummies for G3's birth year.

[4] G2 spousal characteristics include five dummies for spousal education and five age dummies.

Table 1.5 G2's Probability of Being in the Sample as a Function of LBW (or IUGR)

	G2 females (1)	G2 males (2)
G2 birth weight (per kilogram)	0.0028*	0.0071***
	(0.002)	(0.001)
G2 LBW	-0.0126***	-0.0122***
	(0.003)	(0.003)
G2 SGA (5^{th} pctl.)	-0.0074***	-0.0119***
	(0.003)	(0.002)
G2 2SD < mean	-0.0116***	-0.0122***
	(0.004)	(0.003)
Sample size	1,423,811	1,527,356
G2 FT LBW	-0.0137***	-0.0116***
	(0.004)	(0.003)
Sample size	1,385,033	1,484,044

Notes: Standard errors are clustered at child's hospital and year level. Each column of each row is a separate regression. The probability of being observed in our samples is the probability that the singleton G2 is observed giving birth to a singleton (G3) in Taiwan between 1999 and 2006. The estimation sample includes all G2 female (column(1)) or G2 male (column (2)) singleton births in Taiwan between 1978 and 1985 with gestation between 31 and 45 weeks and birth weight between 400 and 6,500 grams. All regressions are based on the most inclusive specification that includes grandmother fixed effects as well as all variables listed in footnotes 2-4 in Tables 1.3 and 1.4. *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 1.6 Effect of Mother's LBW (or IUGR) on Child's LBW (or IUGR)--Adjusted by the Probability of Being Observed

	(1)	(2)	(3)	(4)	(5)
Impact of G2 LBW on G3 LBW	0.0228***	0.0199**	0.0230***	0.0229***	0.0222***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
p-value of t-test: coef. = col. (1)		0.766	0.797	0.801	0.801
Impact of G2 SGA (5th pctl.) on G3 SGA (5th pctl.)	0.0225***	0.0187**	0.0225***	0.0225***	0.0212***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
p-value of t-test: coef. = col. (1)		0.610	0.999	0.999	0.851
Impact of G2 2SD < mean on G3 2SD < mean	0.0192***	0.0148**	0.0192***	0.0192***	0.0182***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
p-value of t-test: coef. = col. (1)		0.540	0.998	0.995	0.886
Sample size	280,030	280,030	280,030	280,030	280,030
Impact of G2 FT LBW on G3 FT LBW	0.0219***	0.0173*	0.0220***	0.0220***	0.0209**
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
p-value of t-test: coef. = col. (1)		0.608	0.997	0.990	0.906
Sample size	255,100	255,100	255,100	255,100	255,100

Notes: Standard errors are clustered at child's hospital-year level. Each column of each row is a separate regression. Column (1) displays fixed-effect estimates from column (6) of Table 1.3; column (2) presents estimates from weighted regression using inverse of the estimated probability of being observed in the maternal sample; column (3) reports estimates from the model controlling estimated probability of observation, $\hat{p}(x)$; column (4) provides estimates from the model controlling $\hat{p}(x)$ and $\hat{p}(x)^2$, a more flexible functional form of the estimated probability of observation; and column (5) reports estimates from the model further including the interaction between mother's LBW (or IUGR) and demeaned predicted probability of observation ($\hat{p}(x)$- $\overline{\hat{p}(x)}$). P-values of t-tests are reported for the equality between each of the estimates from column (2)-(5) and column (1). *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 1.7 Effect of Father's LBW (or IUGR) on Child's LBW (or IUGR)--Adjusted by the Probability of Being Observed

	(1)	(2)	(3)	(4)	(5)
Impact of G2 LBW on G3 LBW	-0.0123	-0.0110	-0.0123	-0.0126	-0.0122
	(0.022)	(0.024)	(0.022)	(0.022)	(0.022)
p-value of t-test: coef. = col. (1)		0.956	0.998	0.992	0.995
Impact of G2 SGA (5th pctl.) on G3 SGA (5th pctl.)	0.0093	0.0095	0.0094	0.0091	0.0090
	(0.020)	(0.022)	(0.020)	(0.020)	(0.020)
p-value of t-test: coef. = col. (1)		0.993	0.997	0.992	0.985
Impact of G2 2SD < mean on G3 2SD < mean	0.0086	0.0082	0.0087	0.0086	0.0083
	(0.019)	(0.021)	(0.019)	(0.019)	(0.019)
p-value of t-test: coef. = col. (1)		0.985	0.995	0.997	0.990
Sample size	125,078	125,078	125,078	125,078	125,078
Impact of G2 FT LBW on G3 FT LBW	0.0294	0.0341	0.0293	0.0286	0.0286
	(0.024)	(0.026)	(0.024)	(0.024)	(0.024)
p-value of t-test: coef. = col. (1)		0.857	0.995	0.972	0.971
Sample size	113,639	113,639	113,639	113,639	113,639

Notes: Standard errors are clustered at child's hospital-year level. Each column of each row is a separate regression. Column (1) displays fixed-effect estimates from column (6) of Table 1.4; column (2) presents estimates from weighted regression using inverse of the estimated probability of being observed in the paternal sample; column (3) reports estimates from the model controlling estimated probability of observation, $\hat{p}(x)$; column (4) provides estimates from the model controlling $\hat{p}(x)$ and $\hat{p}(x)^2$, a more flexible functional form of the estimated probability of observation; and column (5) reports estimates from the model further including the interaction between father's LBW (or IUGR) and demeaned predicted probability of observation ($\hat{p}(x)$- $\overline{\hat{p}(x)}$). P-values of t-tests are reported for the equality between each of the estimates from column (2)-(5) and column (1). *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 1.8 Effect of Mother's LBW (or IUGR) on Child's LBW (or IUGR) by Gender

	Dependent Variables			
	G3 LBW	G3 SGA (5^{th} pctl.)	G3 2SD < mean	G3 FT LBW
	(1)	(2)	(3)	(4)
G2 LBW (or IUGR)	0.0263***	0.0383***	0.0246***	0.0304***
	(0.009)	(0.008)	(0.007)	(0.009)
G2 LBW (or IUGR) x male	-0.0068	-0.0300***	-0.0106*	-0.0164**
	(0.007)	(0.006)	(0.006)	(0.007)
Sample size	280,030	280,030	280,030	255,100

Notes: Standard errors are clustered at child's hospital and year level. Each column is a separate regression. The dependent variables are dummies for child's LBW (or IUGR). Coefficients are reported for dummies for mother's LBW (or corresponding IUGR indicator) and interactions between the dummies and an indicator equal to one if the child is a male. All regressions include the full set of control variables and grandmother fixed effects. *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table 1.9 Effect of Mother's LBW (or IUGR) on Child's LBW (or IUGR) by Gender and SES Groups

	Dependent variables			
	G3 LBW	G3 SGA (5^{th} pctl.)	G3 2SD < mean	G3 FT LBW
	(1)	(2)	(3)	(4)
Panel A: county-level unemployment rate at child's birth year				
(a) G2 LBW (or IUGR) × low unemployment rate × male	0.0170	0.0149	0.0314***	0.0292**
	(0.014)	(0.012)	(0.012)	(0.015)
(b) G2 LBW (or IUGR) × male	-0.0153	-0.0373***	-0.0261***	-0.0310***
	(0.010)	(0.008)	(0.008)	(0.010)
(c) G2 LBW (or IUGR) × low unemployment rate	-0.0241**	-0.0030	-0.0226***	-0.0250**
	(0.010)	(0.008)	(0.008)	(0.010)
(d) G2 LBW (or IUGR)	0.0391***	0.0398***	0.0362***	0.0435***
	(0.010)	(0.009)	(0.009)	(0.011)
(e) row (a) + row (c)	-0.0071	0.0119	0.0088	0.0042
(f) *p-value* for t-test: (a)+(c)=0	0.465	0.149	0.278	0.678
Panel B: town-level average income at child's birth year				
(a) G2 LBW (or IUGR) × high income × male	0.0056	-0.0150	0.0123	0.0175
	(0.014)	(0.012)	(0.012)	(0.015)
(b) G2 LBW (or IUGR) × male	-0.0088	-0.0244***	-0.0149**	-0.0227**
	(0.009)	(0.008)	(0.007)	(0.009)
(c) G2 LBW (or IUGR) × high income	-0.0139	0.0177	-0.0336***	-0.0132
	(0.014)	(0.012)	(0.012)	(0.015)
(d) G2 LBW (or IUGR)	0.0314***	0.0317***	0.0363***	0.0352***
	(0.010)	(0.009)	(0.008)	(0.011)
(e) row (a) + row (c)	-0.0083	0.0027	-0.0213	0.0043
(f) *p-value* for t-test: (a)+(c)=0	0.539	0.813	0.063	0.762

Panel C: county-level average parental education at child's birth year				
(a) G2 LBW (or IUGR) × high education × male	-0.0034	-0.0043	0.0038	0.0247
	(0.014)	(0.012)	(0.012)	(0.015)
(b) G2 LBW (or IUGR) × male	-0.0053	-0.0283***	-0.0118	-0.0250***
	(0.009)	(0.008)	(0.007)	(0.009)
(c) G2 LBW (or IUGR) × high education	-0.0077	-0.0102	-0.0065	-0.0287**
	(0.012)	(0.010)	(0.010)	(0.012)
(d) G2 LBW (or IUGR)	0.0291***	0.0420***	0.0270***	0.0409***
	(0.010)	(0.009)	(0.008)	(0.010)
(e) row (a) + row (c)	-0.0111	-0.0145	-0.0027	-0.0040
(f) *p-value* for t-test: (a)+(c)=0	0.331	0.143	0.776	0.740
Panel D: county-level average educational improvement from mother's birth to child's birth				
(a) G2's LBW (or IUGR) × greater improvement × male	0.0163	0.0047	-0.0062	0.0138
	(0.014)	(0.012)	(0.012)	(0.015)
(b) G2's LBW (or IUGR) × male	-0.0133	-0.0321***	-0.0076	-0.0230**
	(0.009)	(0.008)	(0.008)	(0.010)
(c) G2's LBW (or IUGR) × greater improvement	-0.0132	0.0150	0.0065	-0.0185
	(0.016)	(0.010)	(0.010)	(0.012)
(d) G2's LBW (or IUGR)	0.0316***	0.0313***	0.0216**	0.0389***
	(0.011)	(0.009)	(0.009)	(0.011)
(e) row (a) + row (c)	0.0031	0.0197	0.0003	-0.0047
(f) *p-value* for t-test: (a)+(c)=0	0.761	0.044	0.974	0.695
Sample size	280,030	280,030	280,030	255,100

Notes: Standard errors clustered at child's hospital and year level are reported in parentheses. Each column of each panel is a separate regression. Besides reported variables, all regressions include dummies for high SES groups, interactions between high SES groups and indicators for G3 male birth, the full set of control variables, and grandmother fixed effects. In panel A, low unemployment is a dummy equal to one if the unemployment rate at mother's county of residence at child's birth is below the mean; in panel B, high income is a dummy equal to one if the average town-level income at child's birth is below the mean; in panel C, high education is a dummy equal to one if the percentage of parents'

years of schooling over 12 at mother's county of residence at child's birth is above the mean; in panel D, greater improvement is a dummy equal to one if the difference in county level percentage of greater than 9 years of schooling from G1 to G2 is above the mean. *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

2. Pseudoscience Conspiracy Dies Hard: Evidence from the MMR-Autism Controversy in the United States 1998-2011

2.1. Introduction

It is a prevailing phenomenon that pseudoscientific conspiracies are always found surrounded by their die-hard fans. People form biased beliefs and persistently support these flawed theories by correlating their actual health behaviors without referring to hard proofs. For example, alternative medicine continually gains its popularity without evidence gathered using the scientific method. The cost of such misinformation may lead to delay in treatment or even death. Nevertheless, the mechanism underlying biased beliefs and possible ways to properly address misinformation is not fully understood. We consider the most recent vaccine scare, the MMR-autism controversy, to examine how biased beliefs drive the persistent MMR non-uptake rate over the years when information is mixed.

The MMR-autism controversy was first provoked by a study (Wakefield et al. 1998) published in Lancet. The study links childhood vaccine of Measles-Mumps-Rubella (MMR) to autism. Follow-up studies (Peltola et al. 1998; Farrington et al. 2001; Taylor et al. 2002), authorities (i.e. FDA, and CDC), vaccine manufactures all dispute such a link. This debate has been widely publicized in mass media and has attracted public attention since 2000.[1] Later, the initial study was partially retracted in 2004 and fully retracted in 2010 due to scientific misconduct.[2] However, parents still hold strong skepticism against the vaccine even after the

[1] For example, news titled "house panel asks for study of a vaccine" was published in The New York Times (New York) on April 7, 2000; news titled "state's autism cases continue to increase; little is known why, however one theory on a link to child vaccinations stirs an international feud" was published in Contra Costa Times (California) on April 21, 2000.

[2] After a four-month investigation in 2004, Sunday Times reporter Brain Deer found that the author did not disclose the fact that the research was founded through parents seeking evidence against vaccine manufacturers.

retraction of the paper.[3] Such distrust was further fueled by the "Hannah Poling case" in 2009, which was the first case related to autism to have been awarded compensated by the vaccine court. Even after the partial retraction of the initial paper, the MMR uptake rate declined persistently, resulting in outbreaks of vaccine preventable diseases such as measles. To our surprise, such decline in MMR uptake rate was mainly driven by children of well educated and high-income parents, as documented in the newspaper[4] as well as in previous studies (Wright and Polack 2005; Anderberg et al. 2011).

Studies have shown that media coverage intensity of health information contributes to notable changes in preventive behaviors (Katherine and Brain 2005; Stryker 2003; Yanovitzky and Stryker 2001) and widened education gradient (Anna and Laura 2010). The well-accepted explanation is that the more educated possess better understanding of health risks, thus they react faster. However, there are no disagreements in the attitudes of the information regarding to safety and effectiveness examined by these studies. Indeed, it is important to study how people respond to information when there is ambiguity. As pointed out by psychological literature (Lord, Ross and Lepper 1979; Darley and Gross 1983; Keren 1987; Griffin and Tversky 1992), people may suffer confirmation bias when processing information by misinterpreting ambiguous evidence as confirming his current hypothesis of the world, causing the impact of information to be quite different from the existing literature when information is mixed.

The investigation of the MMR-controversy contributes to the literature with a case where people are exposed to great amount of information with various contents and to contradictory attitudes from various sources. The case helps us to identify differential responses to specific

[3] "Much of the current anti-vaccine movement bases its arguments not on real statistics but on anecdotes, which are powerful, emotional and personal" published in Brattleboro Reformer in 2011, which is the third largest daily newspaper in the state of Vermont.

[4] For example, the article titled "Rich, educated and stupid parents are driving the vaccination crisis" was published in Los Angeles Times on Sep. 3, 2014.

information source and content by education level, which help us understand the joint impacts of information and education on health behaviors when information is mixed. Although studies focusing on this controversy find faster reduction in uptake rate among children of more educated parents (Anderberg et al. 2011) and limited influence of mainstream media on immunization (Smith et al. 2008), no studies provide adequate explanation on the persistent and ever strengthened trend to opt-out the MMR vaccine after the Wakefield paper was retracted. Our paper fills this gap. In particular, we study the differential mechanism that underlies information processing by education level as an explanation for the persistent vaccine scare.

In this paper, we combine data on individual-level immunization records with state-level information exposures to investigate the differential impacts of information on health decisions by the mother's education level. From the National Immunization Survey (NIS) during the period 1998-2011, we obtain vaccination records and demographics for children aged 19-to-35 months. We also assemble state-level information exposures, including passive and active ones. For passive exposures, we obtain relevant disease prevalence rates, from the Office of Special Education Programs and various issues of Morbidity and Mortality Weekly Report, and news counts from LexisNexis Academic database. For active exposures, we collect online search intensity from Google Trends. These exposures are further categorized into groups based on their attitudes, contents, and sources, which helps identify heterogeneous responses to certain features of information. By exploiting variations in differential responses to information by mother's education, we find that strong biased beliefs among mothers with college education are responsible for the persistent non-uptake rate of MMR. Such beliefs cause asymmetric responses to new information based on its attitude and content, which in turn intensify the strength of the beliefs. A one-standard-deviation increase in exposure to information indicating that vaccine

may not be safe leads to 40% increase in the belief coefficient. In contrast, increase in exposure to information encouraging for immunization generally results in a small and insignificant impact on MMR non-uptake rate. Moreover, although authorities' words in the newspaper help to decrease the non-uptake rate, overall impact of main stream media is limited. In contrast, online search results are more influential to mothers when making immunization decisions for their children.

Our results provide empirical evidence on confirmation bias. It is important to keep prudent when conveying information to the public. The impact of an initial misinformation may take a long time to fully address because people may suffer from such bias when processing new information. Moreover, the study also provides implications on how to efficiently and effectively communicate public policy, research results, or even science to the public. By targeting a large and more interested audience, web is a more effective medium than traditional newspaper for spreading the opinions of the authorities.

In the rest of the paper, section 2.2 provides the background of MMR controversy and a literature review. Section 2.3 presents a conceptual framework. Section 2.4 describes how data is collected on individual-level demographics and state-level information exposures. Section 2.5 establishes our empirical model. Section 2.6 and 2.7 provides results and robustness checks. Section 2.8 concludes the study.

2.2. Background

2.2.1. MMR-Autism Controversy

The controversy of MMR-autism is initiated by a paper (Wakefield et al. 1998) published in The Lancet in February of 1998. The article asserts that the measles virus is associated with an inflammatory bowel disease found in autistic children, which is proposed as evidence for the link

between MMR shots and autism risks. However, research using various approaches, larger samples, and a longer study window from different countries all dispute such a link (Taylor et al. 1999; Madsen et al. 2002; DeStefano et al. 2004; Richler et al. 2006). In 2004, the Immunization Safety Review Committee published a final report after examining the scientific evidence and they rejected the link. Due to scientific misconduct, the initial Wakefield study was partially retracted in 2004 and fully retracted in 2010.

In addition to the MMR controversy, two other parallel but related debates exist on the safety of childhood vaccine, which help the MMR controversy escalate into a general vaccine scare and affect the MMR uptake rate. The first debate is the proposed mercury-autism link. In 2001, a published study (Bernard et al. 2001) hypothesizes a link between autism and thimerosal, a preservative used in vaccines. Similarly, later studies do not find sufficient evidence to support such an association (Stehr-Green et al. 2003; Verstraeten et al. 2003; Price et al. 2010).[5] Although the MMR vaccine have never contained thimerosal,[6] the mercury-autism controversy affect uptake rate of MMR indirectly because news generally report the two hypotheses simply as vaccine-autism link without differentiating between them. The second is the criticisms on the heavy vaccine schedule. Some argue that too many vaccines overwhelm the child's immune system, throwing the whole childhood immunization schedule under question, which also has been disputed now by scientific evidence (DeStefano, Price, and Weintraub 2013).

Therefore, during the study period, parents question the safety of childhood vaccines in general, with MMR being the most controversial. In this study, we solely focus on the MMR uptake rate for two reasons. First, the MMR-autism controversy has a more clear and distinct

[5] As a precaution, the Food and Drug Administration (FDA) removed thimerosal from all childhood vaccines except for a few influenza and hepatitis vaccines since 2001 in the United States.
[6] According to Centers for Disease Control and Prevention (CDC), varicella (chickenpox), inactivated polio (IPV), and pneumococcal conjugate vaccines have also never contained thimerosal.

timeline of events with respect to the publication and retraction of the Wakefield paper. Second, a comparison with the other childhood vaccines is not reasonable because most of them are under question due to the mercury-autism link.

Figure 2.1 shows the non-uptake rate, defined as delayed MMR shots, from 1998 to 2011. The annual estimates are obtained using the National Immunization Survey (NIS). Events directly related to the debate on MMR-autism link are labeled. As expected, the trend experienced an overturn during 1998 to 2004 with a local maximum point in 2000 following the Wakefield paper. Interestingly, even after the partial retraction of the initial paper, the non-uptake rate continues to increase, despite mounting scientific evidence that rejects the MMR-autism link and despite claims from health professionals, including FDA, the American Academy of Pediatrics (AAP), Public Health Service, and CDC, that childhood vaccines are safe.

Indeed, after 2004, parents are also more exposed to controversial information filled with emotional personal stories, some of which are even advocated by influential celebrities,[7] which trump the impact of scientific studies.[8] On the one hand, health professionals emphasize the importance of immunization in response to measles outbreaks in U.S. due to low vaccination rate. On the other hand, in March 2008, the vaccine court made the first compensation decision on an autism claim, the Hannah Poling case, which fueled the fear among skeptical parents and was considered by some as government concession on the vaccine-autism link by some.[9] The case

[7] In June 2005, an environmental lawyer and political activist, Robert F. Kennedy Jr., wrote an article titled "Deadly Immunity" in Rolling Stone Magazine, claiming that vaccine is a government/Big Pharma conspiracy. In 2007 and 2009, actress Jenny McCarthy went on the Oprah Winfrey show to promote her new books "Louder than Words: A Mother's Journey in Healing Autism" and "Mother Warriors: A Nation of Parents Healing Autism against All Odds". She also promoted the view of possible autism-vaccine link at the same time.

[8] The New York Times published a news titled "Vaccination: A Hot Debate Still Burning" on April 2010.

[9] Hannah Poling received five vaccines including MMR when she was 19 months old and court concluded that the vaccines worsened a rear and pre-existing cell disorder, resulting in developmental disorders of the child.

attracted wide media coverage and the attention of parents widely under media coverage.[10] On the one hand, actress Amanda Peet teamed with health officials to defend vaccines in December 2008 and the vaccine court ruled against vaccine-autism claims in February 2009. However, on the other hand, two counter-vaccine articles written by Actress Holly Robinson Peete and Actor Jim Carrey[11] were published in March and April of 2009. Over the past decade, people are exposed to mixed information with contradicting attitudes regarding the MMR vaccine with various contents from different sources. The MMR-autism controversy enables us to identify the impact of a specific feature of the information on immunization decision.

Empirical studies that directly examine the impact of MMR-autism controversy on vaccine uptake rates generally focus on a period before 2004, when scientific evidence reached a definite consensus. Using data in the U.S., Smith et al. (2008) posit that the influence of mainstream media on MMR immunization is limited by comparing temporal correlation between MMR non-uptake rate and newspaper coverage. Employing data from the U.K., Anderberg et al. (2011) find that the uptake rate of MMR declined faster in areas where a larger fraction of parents had stayed in education past the age of 18 than in areas with less educated parents. However, both of the studies fail to explain the trend of declining MMR vaccine use after 2004. According to the newspaper, such trend is driven by well-educated parents, the mechanism of which is not examined in previous studies. In order to answer this question, we use a longer study window, which starts from the very first year of the debate till a year after the initial paper was fully retracted. We focus on differential responses in immunization decisions for their children by parental education level when information is mixed.

[10] For example, Akron Beacon Journal (Ohio) published an editorial titled "Rare Conditions" on this case.
[11] In March 2009, actress Holly Robinson Peete wrote an article in Essence magazine to refute comments from actress Amanda Peet and argue for the possible link between vaccines and autism. In April 2009, actor Jim Carrey wrote an editorial to Huffington Post calling parents to be cautious toward the official claims regarding to lack of evidence in vaccine-autism link.

2.2.2. Confirmation Bias

Confirmation bias is a tendency to process information in a way that is consistent to ones prior beliefs. This has long been documented in psychological and cognitive research (Nikerson 1998; Kunda 1999). There are three major manifestations, which contribute to overconfidence and self-perpetuated false beliefs. First, people may selectively search for information in order to prove their pre-existing hypothesis of the world is correct (Kayhan 2013). Second, people may subjectively interpret new evidence based on their beliefs. They value more any information that conforms to their priors and devalue or even ignore information that contradicts their hypothesis. Even provided with the same ambiguous information, beliefs further diverge among people with different initial beliefs (Lord, Ross, and Lepper 1979). Third, people may suffer from biased memory, even if they seek for and evaluate information neutrally. Results from experimental studies are mixed: some suggest that people may recall information that match with their expectations more easily; in contrast, some also posit that unexpected information is more memorable (Oswald and Grosjean 2004).

The MMR-autism controversy provides an appropriate context to examine the impact of the confirmation bias. People adopt a biased view when examining information they received. For example, they misinterpret the Hannah Poling case as recognition that vaccine is not safe and they ignore the findings from scientific research and respond more to celebrity voice, which manifest that they subjectively value information. According to the statistics published by CDC, autism prevalence rate increases from 1 in 150 children in 2000 to 1 in 68 children in 2010. Without clear and acceptable alternative answers to such drastic increase, parents keep holding vaccine as the crucial reason and even throw the government into a trust crisis (Hilton 2007).[12] It

[12] For example, "Haley, the retired University Kentucky professor, put this way: the people saying there is no connection are the ones who caused the problem by making all these vaccines mandatory. The CDC

suggests that parents search for information to test their priors in a one-sided way. More importantly, the larger increase in MMR non-uptake rate among children of well-educated parents as observed in the previous study can be attributed to the possibility that they suffer more confirmation bias. People with high education are more aware of the adverse consequences of autism, and generally consider measles as curable and less harmful. Fear of danger leads people to search for evidence in a biased way and result in exaggerated focus on danger and threat (Friedrick 1993; Gilbert 1998). Highly educated people may also have higher perceived knowledge, which may not be necessarily consistent with their actual knowledge level. Using a field experiment, Park et al. (2013) find that confirmation bias is more pronounced among investors with higher perceived knowledge about the market. Although previous literature focus on differential confirmation bias by anxiety level (Remmerswaal 2014)[13] and gender (Traut-Mattausch 2011), we, in particular, examine whether variations in vaccine decisions by parental education level are attributable to different degrees of confirmation bias.

2.3. A Conceptual Framework

In the Bayesian information processing model, rational people hold common priors and update their beliefs in the same direction in response to a given signal. To incorporate confirmation bias into economic model, previous studies used different ways to revise the traditional assumptions (Rabin and Schrag 1999; Wing 2004). Following the literature, we lay out a simple model of information and vaccination. We assume an extreme case that college (well-educated) parents suffer confirmation bias by valuing information confirms to their priors more, but non-college (less-educated) parents do not. Our objective is to articulate with the conceptual framework what will be observed with our data if differential confirmation bias by education level exists.

---

mad a big mistake, and they don't want to admit it", which is published in the news titled "Family faces uncertainty in dealing with autism" in the Leader-Telegram (Wisconsin).

[13] See Matthew and MacLeod (1994) for a review.

2.3.1. The Set-Up

There are two states of the world, $\Omega \in \{H, S\}$, where H denotes the state in which vaccine is harmful and S denotes the state in which vaccine is safe. The common priors are that both states are equally likely. Parents are making vaccine decisions on behalf of their children. We assume that parents are altruistic toward their children and fully internalize their children's health benefits and costs of vaccine. The two possible actions parents take are $a \in \{v, r\}$, where v denotes taking vaccine and r denotes refusing to take vaccine.

We assume that the health benefits and costs of vaccine are heterogeneous among families. To a family the health benefit of vaccine is, in utility terms, $\alpha \geq 0$. And in the case that vaccine is harmful, the health cost of vaccine is $\beta \geq \alpha$. We thus have $U(S, v) = \alpha$, $U(H, v) = \alpha - \beta$, and $U(H, r) = U(S, r) = 0$.

We assume that there are two groups of parents, college educated and non-college educated. The difference between them pertains to how they process information, which will be covered below. Within each group, each family has a different health benefits and costs of vaccine. Each family is identified with the ratio $\frac{\alpha}{\beta}$. There is a unit mass of college educated parents with $\frac{\alpha}{\beta}$ distributed on [0,1] according to distribution F_c; there is another unit mass of non-college educated parents with $\frac{\alpha}{\beta}$ distributed on [0,1] according to F_{nc}. For simplicity, we assume that $F_c = F_{nc} = F$.

In every period $t \in \{1, 2\}$, parents receive a signal $s \in \{h, s\}$ that is correlated with the true state of the world. Signals received at different t are independently distributed, with $\Pr(h|H) = \Pr(s|S) \in \{\frac{1}{2}, 1\}$. The signal can be interpreted as the information regarding whether vaccine is

harmful or safe. In this context, h is the information that vaccine is harmful and s is the information that vaccine is safe. We will adopt this interpretation from now on.

We assume that the two groups of parents perceive the values of $\Pr(h|H) = \Pr(s|S)$ differently. For non-college educated parents, they consider $\Pr(h|H) = \Pr(s|S) = p$. For college educated parents, they perceive $\Pr(h|H) = \Pr(s|S) = p$ only when they consider H and S are equally likely. If they consider that H is more likely, college educated parents perceive that $\Pr(h|H) = \Pr(s|S) = p' > p$ when h is received and that $\Pr(h|H) = \Pr(s|S) = p'' < p$ when s is received. If they consider that S is more likely, college educated parents perceive that $\Pr(h|H) = \Pr(s|S) = p' > p$ when s is received and that $\Pr(h|H) = \Pr(s|S) = p'' < p$ when h is received. In other words, college educated parents suffer from some sort of confirmation bias, who consider a new information to come from a more informative source when it conforms to their existing beliefs.

2.3.2. Analysis

Parents choose to vaccine their children if and only if the expected payoff from doing so is greater than zero. In period $\mathrm{t} = 1$ before receiving any new information, this means that when $\frac{\alpha}{\beta} \geq \frac{1}{2}$. The proportions of college and non-college educated parents' taking vaccine are thus $1 - \mathrm{F}(\frac{1}{2})$. Upon receiving a new signal in $t = 1$ that vaccine is harmful, both college and non-college parents update their beliefs according to

$$\mu_1^c(h) = \mu_1^{nc}(h) = \frac{\frac{1}{2}p}{\frac{1}{2}p + \frac{1}{2}(1-p)} = p. \tag{2.1}$$

The corresponding threshold for taking vaccine is $\frac{\alpha}{\beta} \geq p$. The proportions of college and non-college educated parents' taking vaccine are thus $1 - F(p)$. Since $p > \frac{1}{2}$, $1 - F(p) < 1 - F(\frac{1}{2})$.

We summarize this observation with the following two trends we expected to find in our data:

Proposition 2.1. In period $t = 1$, upon receiving a signal that vaccine is harmful, the proportions of college and non-college educated parents' taking vaccine for their children decrease.

Note that p becomes the prior beliefs of both groups of parents at the beginning of $t = 2$. Since $p > \frac{1}{2}$, college and non-college educated parents start to process information differently. Upon receiving a signal in $t = 2$ that vaccine is actually safe, college and non-college educated parents' updated beliefs are

$$\mu_2^c(s) = \frac{p(1-p'')}{p(1-p'')+(1-p)p''},$$

$$\mu_1^{nc}(s) = \frac{p(1-p)}{p(1-p)+(1-p)p}. \tag{2.2}$$

Given that $p'' < p$, $\mu_2^c(s) > \mu_1^{\mathrm{nc}}(s)$, which further implies that $1 - F(p) < 1 - F\left(\mu_2^c(s)\right) < 1 - F(\mu_1^{nc}(s))$.

On the other hand, upon receiving a signal in $t = 2$ that vaccine is indeed harmful, college and non-college educated parents' updated beliefs are

$$\mu_2^c(h) = \frac{pp'}{pp'+(1-p)(1-p')},$$

$$\mu_1^{nc}(h) = \frac{p^2}{p^2+(1-p)^2}. \tag{2.3}$$

Given that $p' > p$, $\mu_2^c(h) > \mu_1^{nc}(h)$, which further implies that $1 - F(\mu_2^c(h) < 1 - F\left(\mu_1^{nc}(h)\right) < 1 - F(p)$. We summarize the above with the following phenomenon we expect to find in our data:

Proposition 2.2. Having received in $t = 1$ a news that vaccine is harmful, upon receiving a news in $t = 2$ that it is actually safe, the proportions of college and non-college educated parents' taking vaccine for their children increase, and the increase is higher for non-college educated parents; upon receiving a news in $t = 2$ that it is indeed harmful, the proportions of college and

non-college educated parents' taking vaccine for their children decrease, and the decrease is higher for college-educated parents.

2.4. Data and Sample

We combined data on individual-level immunization record and state-level information exposures to empirically examine whether the strong and persistent trend in MMR non-uptake rate is driven by biased beliefs by education level when information is mixed. If the answer is confirmed, we further examine the mechanism underlying such biased beliefs. We calculated cumulative exposures for each type of information during our study period to capture general features of the information by state that are available to parents. To be comparable across information sources, contents, and attitudes, we obtained z-scores using the mean and standard deviation of 51 states for a given information exposure.

2.4.1. Individual Immunization Records and State characteristics

We obtained immunization records and demographics for children aged 19-to-35 months from the National Immunization Survey (NIS) 1998-2011, which is an annual telephone survey administered to estimate immunization coverage. Households with children in the target age range are randomly selected each year and asked a series of vaccination and demographics questions. Child's primary care physician was contacted under consent to verify vaccination records. For accuracy, we restricted our sample to children with valid provider data. In the survey year 2011, only landline sample is included in order to be consistent with previous years. We observed up-to-date status of MMR shots, demographics for the child and mother, socioeconomic status of the household, and the child's health care facility type. The data also provides state identifier, which is used to merge state-level information exposures. Conditional non-missing values in demographics, we obtained a total of 271,478 observations.

In order to capture some time-variant state-level characteristics, we obtained percentage of uninsured children under 18 to all people by state-year from Current Population Survey Annual Social and Economic Supplement (CPS ASEC)[14]. We collected percentage of immigration share of residence by state-year from the Yearbook of Immigration Statistics. And we acquired estimated resident population by state-year from Census Bureau.

2.4.2. State-level Information

For information exposures, we consider both passive receptions and active searches. For passive information exposures, we obtain state-level autism prevalence rate, total reported cases for measles, mumps and rubella, and news counts from state news outlets. For active information exposures, we acquire state-level online search intensity on relative topics. Differential impacts of active information exposures on strength of false beliefs by education level not only capture subjective valuation of information but also absorb selectively test for priors under confirmation bias. We further group both types of information based on source, content, and attitude.

We estimated autism prevalence rate by state-year using autism to total counts in special education program for age 6 to 21 from the Office of Special Education Programs (OSEP) during 1998-2011[15]. OSEP maintains standardized compilations of state counts of children receiving free public education services. The counts are classified into 13 primary disability categories defined under the Individual with Disability Education Act (IDEA). Autism is among one of them. We collected total reported cases for indigenous measles, mumps and rubella during 1998-2011 by state-year from various issues of Morbidity and Mortality Weekly Reports (MMWR). The statistics are compiled from reports sent by state health departments and territories to the

[14] Uninsured rate by state in the year 1998 is not available to us. We imputed it by assuming values in the year 1999 are the average of those in the year 1998 and 2000.

[15] We imputed the counts as the average of year 1998 and 2000 for year 1999, due to the data in that year is not available. For the state Vermont, values for year 2007 and year 2008 are missing. We imputed them using the average of 2006 and 2009.

National Notifiable Diseases Surveillance System. In order to make the data comparable across states, we normalized the values using estimated resident population.

For newspaper coverage, we count number of stories captured by the search terms "MMR" and "autism" or "vaccine" and "autism" or "measles" and "autism".[16] We searched LexisNexis Academic from 1998 to 2012. News from the outlet of a given state is used to construct the newspaper coverage for that state. National newspapers were not included because we expect they have the same effect across states. We obtained a total of 208 pieces of news. Within our search, newspaper starts to publicize the controversy in 2000. We find variations in newspaper coverage by state. There are 35 states with at least one piece of news related to the MMR-autism controversy and the 16 states in Table 2.1 accounted for 61% of the total news counts we obtained. We further analyze each piece of news to determine its attitude, positive if encourages vaccination, negative if discourage vaccination or narrative. Based on the content of the news, we assign different indicators if the news includes opinions from parents, words from authorities, and scientific proof.[17] Authorities are defined as the government, health care agencies, health professionals and researchers. We created variables by state-year for total news count, percentage of news with positive attitudes, opinions from parents, words from authorities, and evidence from scientific study.

Google Trends provides us a search index that represents a relative value of search intensity for a term in a given state-year. The index ranges from 0 to 100, a higher value indicating more intensive search in that state-year. The Trends eliminates repeated queries from a certain user

[16] We did not use search term "thimerosal" and "vaccine" because thimerosal-autism link is irrelevant to MMR vaccine. However, if the news only generally mention the relationship between vaccine and autism, we consider it a relevant news count because it potentially affect uptake rate of MMR.

[17] The newspaper coverage is generally filled with parents doubt against vaccine, authorities' persuasion for immunization, and scientific proof against the link with autism. After scrutinizing the articles, only five of them contained parents' opinions encouraging vaccination and only one of them involved suggestions against vaccine by authorities.

over a short period of time and only analyzes data if search volume is over a certain threshold and results are normalized by total number of Google searched done nationwide in that year. We obtained search index at state-year level for terms "Autism (Disease)" "Measles (Disease)", "Mumps (Disease)", "Rubella (Disease)", "vaccine and autism". These search indexes are available to us post 2004, the year Google Trends was put into use. Due to extreme low search volume for the topic "vaccine and autism", the index is available since 2008. As an analogous to disease prevalence rate, we average values for measles, mumps and rubella together to obtain a combined search index for disease outbreaks.

2.4.3. Sample Statistics

Table 2.2 presents sample statistics for individual demographics. Estimates in column (1) and (2) are obtained using the full sample and estimates in column (3) and (4) are acquired from children of college and non-college mothers, respectively. Table 2.3 presents characteristics of the cumulative state-level information exposures. Means and standard deviations are reported for each type of information using the final regression sample. We also predict the effect of each information variable on non-uptake rate of MMR. A positive sign indicates that we expect the type of information discourages vaccination decision and thus is a harmful signal. In contrast, a negative sign indicates that the information encourage immunization and thus is a safe signal.

2.4.4. Differential Responses between College and Non-college Mothers

For each child i in state s in year t, we consider

$$y_{ist} = \boldsymbol{X_{ist}}\beta + \lambda_{st} + \tau_s + \eta_t + \varepsilon_{ist}, \quad (2.4)$$

The outcome of interest is an indicator for none up-to-date shots of MMR. $\boldsymbol{X_{ist}}$ is a vector of observables for the child, λ_{st} is time-variant state-level variables, τ_s is state fixed effects and η_t is year fixed effects. ε_{ist} is the idiosyncratic error term. We include child's gender, firstborn

status, dummies for age group, dummies for race categories, and dummies for facility types, indicator equal to one if the child is moved from a different state, mother's marital status, dummies for mother's education level, and dummies for mother's age group. For time-variant state characteristics, we have percentage of uninsured children under 18 among all uninsured, immigration share of residence, and population by state-year.

Table 2.4 reports estimated coefficients of year dummies in equation (1) for our full sample (in column (1)), and children of college (in column (2)) and non college mothers (in column (3)). In response to newspaper coverage, there is an increase in MMR non-uptake rate among both college and non-college groups in 2000. In response to the partial retraction of Wakefield paper in 2004, MMR non-uptake rate among low education group decreased significantly by 1.62 percentage points compared to that in 1998. In contrast, the vaccine rate is not statistically different from that in 1998 for high education group. We find that post the year 2004, the unexpectedly persistent and strengthened trend to opt out MMR shots is mainly driven by children of college mothers. For the well-educated group, estimates are significantly positive and increase in magnitude after 2004. In contrast, for low education group, estimates are generally small and insignificant, except for the year 2009. Due to the impact of the anti-vaccine voice from celebrities, non-uptake rate increase for both groups in 2009. However, only high education group exhibit a carry-over effect in later years.

Compared to the baseline,[18] MMR non-uptake rate increased by 31.5% in 2000 and remained unchanged in 2004 for college group. In contrast, the non-uptake rate for non-college group increased by 13.5% in 2000 and decreased by 18.6% in 2004. Though non-uptake rate increased in both groups in response to news coverage, asymmetric responses to information that suggests safety of vaccines in 2004 are only found in college group, which are consistent with the

[18] The baseline non-uptake rate for MMR is 0.06082 for college group and 0.08714 for non-college group.

predictions of our conceptual framework under the assumption that college parents suffer from confirmation bias more. The top graph in Figure 2.2 plots the estimated coefficients in column (2) and (3) of Table 2.4. And the bottom graph depicts the differences in the annual estimates between college and non-college groups. Education gradient widened continuously after 1998 and it shows limited impacts of events in following years, suggesting that the strong and persistent trend to delay MMR can be attributable to biased beliefs formed since 1998 among college group.

2.5. Model

To investigate how strong is the biased beliefs that drove the MMR non-uptake rate over years for high education group, we consider

$$y_{ist} = college \times f(year_t) + college \times post + \boldsymbol{X_{ist}}\beta + \lambda_{st} + \tau_s + \eta_t + \varepsilon_{ist}, \quad (2.5)$$

where $f(year_t)$ is a time trend, college is a dummy equal to one for child of mother with at least a bachelor degree, and $post$ is a dummy equal to one for years after 2004. All the other controls are the same as those in equation (2.5). The term, $college \times f(year_t)$, captures the impact of biased beliefs formed by college parents since 1998. And the term, $college \times post$, identifies any deviations from the trend following the retraction of the Wakefield paper in 2004. We change the definition of the $post$ dummy to identify the impact of different events over the years.

To further examine the underlying mechanism for the persistence of the biased belief, we consider

$$y_{ist} = college \times f(year_t) + college \times f(year_t) \times info_s + college \times post + \boldsymbol{X_{ist}}\beta + \lambda_{st} + \tau_s + \eta_t + \varepsilon_{ist}, \quad (2.6)$$

where $info_s$ is a vector for state-level cumulative information exposures, which are used as proxy for the overall features of information available to parents over time. All the other controls

are the same as those in equation (2.6). The three layer interaction term, $college$ x $f(year_t)$ x $info_s$, identifies the effect of exposures to each specific information type on biased beliefs for high education group.

We choose log linear as the functional form for the time trend. It captures an increasing average impact of beliefs on immunization decision with a diminishing marginal effect over years, which is also consistent with the plotted differences between estimates from college and non-college samples depicted in Figure 2.2.

2.6. Results

2.6.1. The Impact of Biased Beliefs

Table 2.5 presents results from equation (2.5). In the odd columns, we control for a post dummy equal to one for years after 2004. In the even columns, we include indicators for each single year after 2004 to capture any deviations from the year trend due to events in the years following the partial retraction of the Wakefield paper. In columns (1), we do not include area specific year effect; in columns (2), we add $region$ x $year$ fixed effects; in column (3), we control for $division$ x $year$ fixed effects;[19] and in columns (4), we include state fixed effects x log year trend. The results are robust after considering area specific year trend.

The strong and significant estimates for year trend and college interaction suggest that college mothers hold stronger beliefs against MMR vaccine over years compared to non-college mothers. Estimates for the post dummies for 2004 are small and insignificant, indicating that the partial retraction of the Wakefield paper generally does not have meaningful impact on the immunization decisions of college mothers. Examining the effects from each single year after 2004 also supports for the finding. Except for 2010, we detect no deviations from the year trend.

[19] We group states into 9 divisions based on U.S. Census Bureau.

Indeed, the significant estimates for the indicator of year 2010 are consistent with our hypothesis that college mothers suffer more confirmation bias. MMR non-uptake rate increase in both groups in response to anti-vaccine voice form celebrities and aftershock of the Hannah Poling's case. However, the impact linger into following years only in college group, resulting in significant difference observed in 2010 when compared to non-college group. Therefore, the persistent increasing non-uptake rate of MMR is mainly driven by the biased beliefs of college parents.

2.6.2. Mechanisms for Biased Beliefs

Table 2.6 presents estimates from equation (2.6) and is organized similar to Table 2.5. Each type of information exposures described in Table 2.3 is interacted with the college indicator and a log linear year trend. We report z-scores for information variables because standardized variables are comparable and provide meaningful interpretation. Results are similar using raw values of these information exposures as presented in Table C1 in Appendix. The findings are robust after including different area specific time trends.

In all specifications, coefficients of $college \times post$ are not materially affect by the inclusion of information interactions. We find significant impact on the belief coefficient for four types of specific information exposure: reported total cases of measles, mumps and rubella, percentage of news with words from authorities in newspaper, online search intensity for disease outbreaks, and online search intensity for the topic "vaccine and autism". A one-standard-deviation increase in reported cases of diseases leads to a decrease of 13.3-15.3% in the belief coefficient. A one-standard-deviation increase in percentage of news with words from authorities is associated with a decline of 24.6-28.1% in the belief coefficient. A one-standard-deviation increase in web search index for vaccine preventable disease is associated with a decline of 28.9-32.8% drop in

the belief coefficient. In contrast, a one-standard-deviation increase in web search index for "vaccine and autism" is related with 22.2-26.3% increase in the belief coefficient. The impacts of these information measures are consistent with our expectations. Interestingly, the impact of autism prevalence rate is negligible and insignificant. Though we only use a crude measure, it suggests that mothers' fear for adverse consequence from vaccine is widespread regardless of the actual autism prevalence rate.

We further examine whether college parents value information from various sources differently. We focus on three information sources: disease prevalence rate, newspaper, and web searches. We create composite information measures according to the expected sign of each single measure by averaging the z-scores of information from the same source. Information is treated as a harmful signal if its expected impact on MMR non-uptake rate is positive. Likewise, information is regarded as a safe signal if its expected sign is negative. The impact of news count is inconclusive. More media attention may help the mass to understand the origin and consequence of the controversy, which lead to a negative impact on MMR non-uptake rate. But it may result in a positive impact on the MMR non-uptake rate by providing more chance for biased readers to selectively test their hypothesis especially when most of the news presents a mix of scientific evidence, words from authorities, and personal stories. Luckily, the impact of news count in Table 2.6 is trivial and insignificant. And we find that grouping it into either a safe (in Table C2) or a harmful signal (in Table 2.7) does not affect our results.

Table 2.7 displays estimates using composite information exposures based on information source and attitude. Although college mothers vaccinate their children in response to words from authorities, the overall impact of newspaper on immunization decision is limited, which is consistent with the literature (Smith et al. 2008). However, a one-standard deviation of safe

signals from disease prevalence rate leads to 14.8-16.1% decrease in the belief coefficient. The impact of safe signals from web searches is similar but only significant at 10 percent level and less robust when considering area specific time trend. In contrast, a one-standard-deviation increase in harmful signals from web searches is associated with 29.9-41.6% increase in the belief coefficient, suggesting a strong and significant impact. Compared to traditional media, web tends out to be more influential to mothers' vaccination decisions. However, under confirmation bias, college mothers may actively search for information to confirm their beliefs in a biased way, leading to more significant impact of harmful signals in general and thus further sustain their priors.

In Table 2.8, we further aggregate information exposures solely based on the expected effects. That is, we obtain harmful composite by averaging z-scores of all information variables with an expected positive sign and news count. And we obtain safe composite by averaging z-scores of all information variables with an expected negative sign. Our results are not affected if we consider news count as a safe signal as presented in Table C3. The findings further confirm that college parents suffer more confirmation bias in our study. A one-standard-deviation increase in the composite harmful signal leads to 43.0-51.8% increase in the belief coefficient. In contrast, a one-standard-deviation increase in the composite safe signal only leads to 14.18-20.14% decrease in the belief coefficient and the impact is insignificant. The results suggest that college parents filter the information they respond to based on their existing and biased priors, further strengthening their false beliefs and lead to strong and persistent MMR non-uptake rate even after the Wakefield paper was partially retracted.

2.7. Robustness Check

We perform two types of robustness check. First, we only include children not moved from a different state. We expect more precise estimated impacts of information exposures because the estimates will be less contaminated by information features in the other states. Second, we exclude parents unable to speak and read English from our sample. Because these parents are immune to most of the information signals we examined in this study even though they live in a state exposed to intensive information during the MMR-autism controversy. To the extent that this factor is correlated with parental education level, our results may be driven by the gap in information availability instead of difference in understanding the same type of information between college and non-college parents. Unfortunately, the NIS sample does not provide us with English speaking status of parents. Therefore, we only include non-Hispanic White and Black children in our sample to check whether our main results still hold with major English speaking population.

Tables 2.9, 2.10, 2.11 and 2.12 report estimates from the sample restricted to children not moved from a different state. We find robust results. Compared to non-college parents, college parents form strong biased beliefs over the years, which affect their immunization decisions for their child. In addition to our main results, we also find that parents opinions in the newspaper leads to an increase in the biased beliefs. In contrast, scientific evidence in the newspaper leads to a decrease in the biased beliefs. However, the estimates are only significant at 10 percent level and the significance goes away when considering area specific time trends. Moreover, results examining impact of information sources suggest a stronger and significant effect of safe signals from web searches, though the estimates are smaller compared to harmful signals from web searches. Our finding that online search results are more influential to mothers in immunization

decisions for their children is further confirmed. Furthermore, similar results are obtained using the sample including only non-Hispanic White and Black children as presented in Tables C4, C5, C6, and C7 in Appendix.

2.8. Conclusion

This study considers the most recent vaccine scare, the MMR-autism controversy, as a platform for studying differential responses to information by education level when information is mixed.

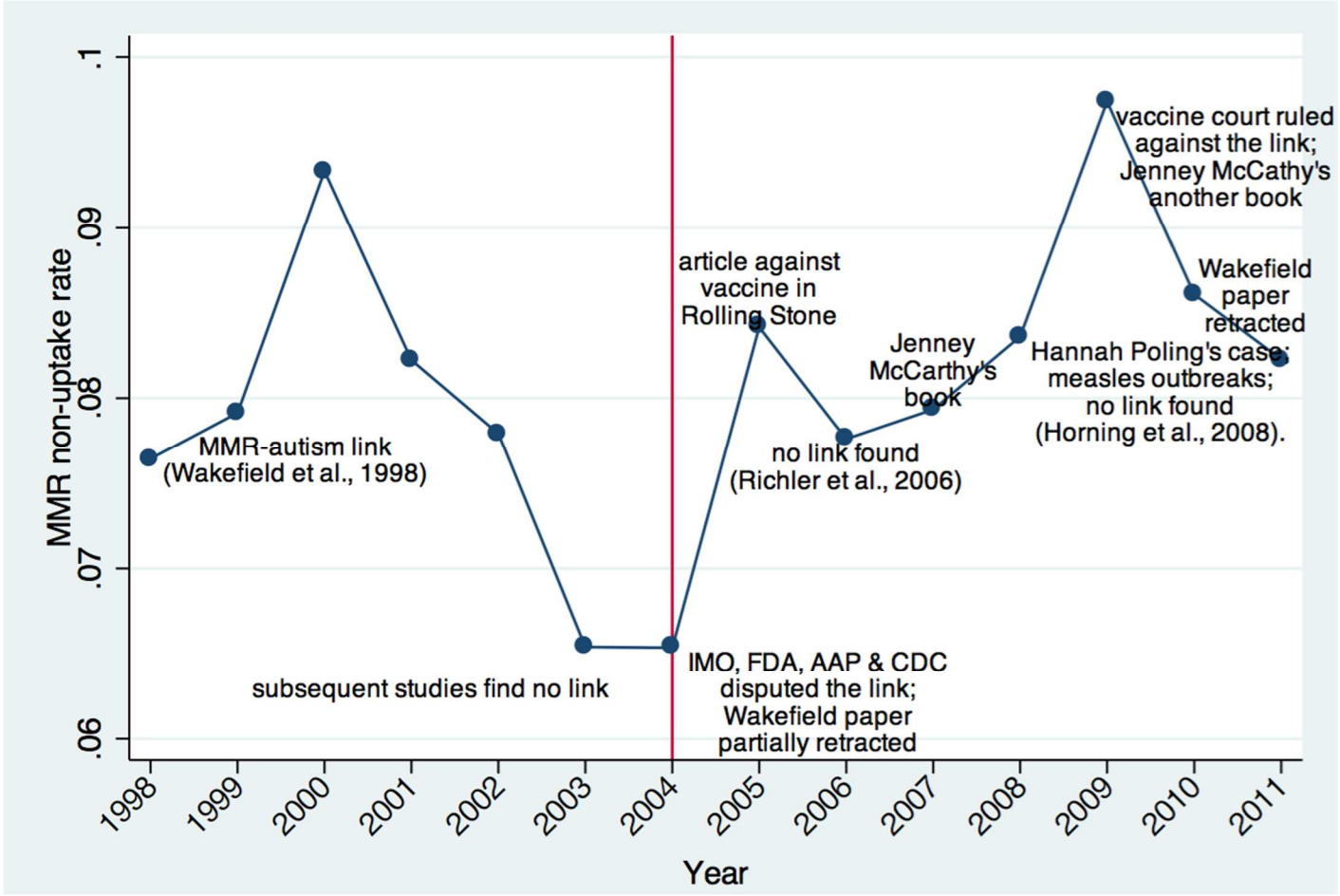
The MMR-autism controversy was first provoked by a paper (Wakefield et al. 1998), which links autism to the childhood vaccine of MMR. Follow-up studies disprove such a link. And the Wakefield paper was partially and then fully retracted in 2004 and 2010. Even after the retraction of the paper, parents still hold strong skepticism against MMR vaccine especially among well-educated mothers. During the wake, people are exposed to tremendous information with contradicting attitudes and mixed contents from various sources. As psychology literature documented that people may suffer confirmatory bias when processing information with ambiguous evidence, we focus on whether such bias beliefs are the driving factors for the persistently increasing MMR non uptake rate among well-educated mothers over the years. We also investigate the underlying mechanism for such biased beliefs.

In order to empirically study the impact of information on health decisions, we combined data on individual-level immunization records and state-level information exposures. We obtained vaccine record and demographics for children aged 19-to-35 months from the National Immunization Survey (NIS) for the years 1998-2011. We assembled both passive and active cumulative information exposures at state-level during the study period. For passive exposures,

we collected relevant disease prevalence rates and newspaper coverage. For active information exposures, we acquired online search index for related topics.

Our results show that the persistently increasing MMR non-uptake trends are driven by biased beliefs among well-educated mothers. Suffering from confirmation bias, well-educated mothers respond more to harmful signals that confirming their priors, further intensifying these false beliefs. A one-standard-deviation increase in harmful signals in general is associated with around 45% increase in belief coefficients. Although college mothers decide to vaccinate their children more in response to words from authorities, the overall impact of newspaper coverage is limited. In contrast, web search results are more influential to the immunization decisions.

Figure 2.1 MMR Non-Uptake Rate and the Timeline of MMR-Autism Controversy



Notes: the figure plots annual estimated percentage of children aged 19-to-35 months without up-to-date shot of MMR as indicated in their immunization record and confirmed by their providers. Estimates are obtained from National Immunization Survey 1998-2011 and weighted to represent the whole population of US, excluding Virgin Island. In the survey year 2011, only landline sample is used in order to be consistent with previous years. The sample size is 280,379, which includes all individuals with valid provider data.

Table 2.1 Newspaper Coverage by State

State	Newspaper	News counts
(1)	(2)	(3)
California	Contra Costa Times	9
California	San Jose Mercury News	7
California	Orange County Register	3
Florida	St. Petersburg Times	6
Florida	Florida Times-Union	4
Minnesota	St. Paul Pioneer Press	5
Missouri	St. Louis Post-Dispatch	3
Nebraska	Lincoln Journal Star	3
Nevada	Las Vegas Review-Journal	3
New York	The New York Times	7
New York	Daily News	5
New York	The New York Post	5
New York	Buffalo News	3
Ohio	Akron Beacon Journal	7
Oklahoma	The Oklahoman	5
Pennsylvania	Pittsburgh Post-Gazette	7
Pennsylvania	Sunday News	4
Pennsylvania	York Daily Record	3
Pennsylvania	Intelligencer Journal / New Era	3
Texas	Austin American-Statesman	4
Texas	El Paso Times	3
Virginia	The Roanoke Times	3
Washington	The Columbian	3
West Virginia	Charleston Gazette	5
Wisconsin	Wisconsin State Journal	7
Wisconsin	The Capital Times	6
Wyoming	Wyoming Tribune-Eagle	3

Notes: News counts are obtained from LexisNexis Academic using search term search terms “MMR” and “autism” or “vaccine” and “autism” or “measles” and “autism”. News from the outlet of a given state is used to construct the newspaper coverage for that state. Only newspapers with at least three counts during 1998-2011 are listed in the table.

Table 2.2 Sample Statistics for Individual Demographics

	Full sample		College		Non college	
	Mean	St.d.	Mean	St.d.	Mean	st.d.
	(1)	(2)	(3)	(4)	(5)	(6)
MMR non-receipt rate	0.076	(0.266)	0.061	(0.239)	0.087	(0.282)
Child's age group (reference: 19-23 months)						
24-29 months	0.352	(0.478)	0.353	(0.478)	0.352	(0.478)
30-35 months	0.351	(0.477)	0.354	(0.478)	0.348	(0.476)
Child's race (reference: Hispanic)						
Non-Hispanic white	0.599	(0.490)	0.743	(0.437)	0.499	(0.500)
Non-Hispanic black	0.125	(0.330)	0.071	(0.257)	0.161	(0.368)
Other non-Hispanic	0.082	(0.275)	0.092	(0.289)	0.076	(0.265)
% of male	0.512	(0.500)	0.512	(0.500)	0.511	(0.500)
% of first born	0.420	(0.493)	0.452	(0.498)	0.077	(0.267)
% of moved from a different state	0.081	(0.274)	0.087	(0.283)	0.397	(0.489)
Mother's education (reference: <12 years)						
% 12 years	0.249	(0.432)	--	--	0.420	(0.494)
% >12 years, non-college	0.220	(0.414)	--	--	0.371	(0.483)
% college graduate	0.408	(0.491)	--	--	--	--
Mother's age group (reference: <19)						
20-29	0.393	(0.488)	0.198	(0.398)	0.527	(0.499)
> 30	0.583	(0.493)	0.801	(0.399)	0.433	(0.495)
% of married	0.739	(0.439)	0.917	(0.276)	0.617	(0.486)
Family income (reference: 0-30K)						
% 30-50K	0.187	(0.390)	0.158	(0.364)	0.207	(0.405)
% >50K	0.425	(0.494)	0.719	(0.449)	0.223	(0.416)
unknown	0.061	(0.239)	0.023	(0.150)	0.087	(0.282)
Facility type (reference: public)						
% private	0.576	(0.494)	0.688	(0.463)	0.499	(0.500)
% others	0.210	(0.407)	0.183	(0.386)	0.229	(0.420)
% mixed	0.085	(0.278)	0.070	(0.255)	0.095	(0.293)
Sample size	271,478		110,688		160,790	

Notes: Sample statistics for individual demographics are reported for the full sample in column (1) and (2), for children of mothers with at least bachelor degree in column (3) and (4), for children of mothers without a bachelor degree in column (5) and (6).

Table 2.3 Cumulative State-level Information Exposures from 1998-2011

	Data source (1)	Data year (2)	Mean (3)	St.d (4)	Predicted effect (5)
Passive information exposures:					
Disease prevalence rate					
Autism	OSEP	1998-2011	0.455	(0.141)	+
Reported cases of measles, mumps and rubella	MMWR	1998-2011	0.636	(1.100)	-
News coverage					
News count	LexisNexis	2000-2011	5.742	(7.541)	+/-
Percent encouraging vaccination	LexisNexis	2000-2011	0.975	(1.107)	-
Percent with opinions from parents	LexisNexis	2000-2011	1.319	(1.554)	+
Percent with opinions from authorities	LexisNexis	2000-2011	1.047	(1.360)	-
Percent with scientific evidence	LexisNexis	2000-2011	1.319	(1.554)	-
Active information exposures:					
Web search intensity					
search index for measles, mumps and rubella	Google Trend	2004-2011	599.396	(549.909)	-
search index for Autism	Google Trend	2004-2011	532.843	(102.853)	+
search index for "vaccine and autism"	Google Trend	2008-2011	31.571	(84.678)	+

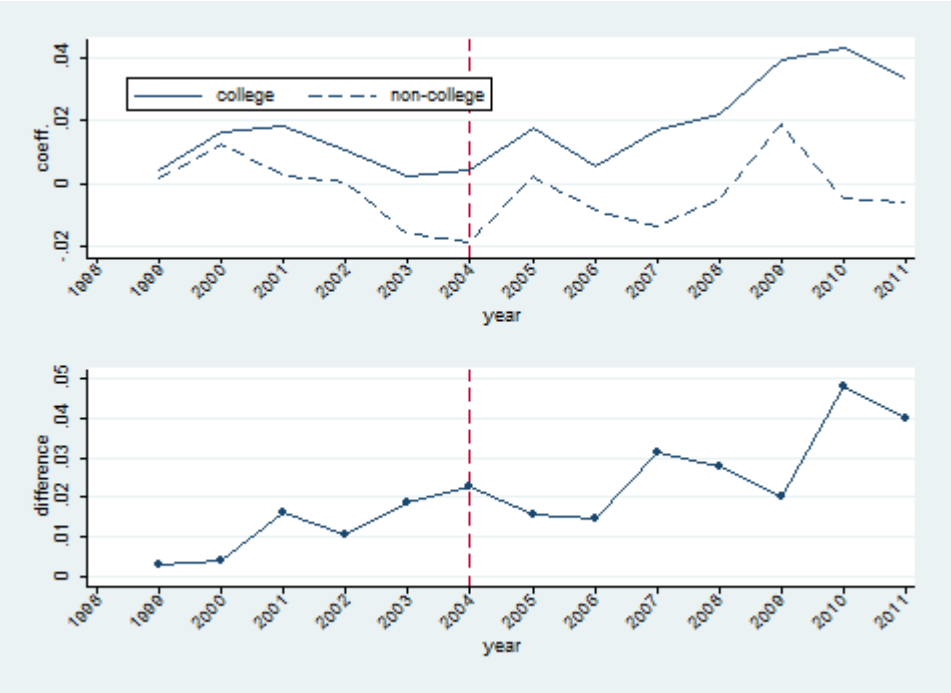
Notes: Each information exposure is the cumulative value by state-year during our study period. Column (1) presents the data source; column (2) lists available data years, column (3) and (4) report sample statistics, and column (5) displays predicted effect on MMR non-uptake rate. A positive sign suggest that the given information exposure is expected to increase the non-uptake rate of MMR. And a negative sign means that the given information exposure is expected to decrease the non-uptake rate of MMR. The sample size used for estimation is 271,478.

Table 2.4 MMR Non-Uptake Rate Compared to 1998

VARIABLES	Full sample (1)	College (2)	Non-college (3)
Year=1999	0.0027	0.0060	0.0016
	(0.004)	(0.006)	(0.006)
Year=2000	0.0136***	0.0192***	0.0118*
	(0.005)	(0.006)	(0.006)
Year=2001	0.0075	0.0209***	0.0028
	(0.005)	(0.006)	(0.006)
Year=2002	0.0045	0.0131**	0.0018
	(0.005)	(0.006)	(0.007)
Year=2003	-0.0094**	0.0011	-0.0135**
	(0.004)	(0.006)	(0.006)
Year=2004	-0.0105**	0.0047	-0.0162***
	(0.004)	(0.006)	(0.006)
Year=2005	0.0086*	0.0197***	0.0047
	(0.005)	(0.007)	(0.007)
Year=2006	-0.0021	0.0088	-0.0058
	(0.005)	(0.007)	(0.007)
Year=2007	-0.0010	0.0198***	-0.0088
	(0.005)	(0.007)	(0.006)
Year=2008	0.0068	0.0272***	-0.0008
	(0.005)	(0.007)	(0.007)
Year=2009	0.0294***	0.0418***	0.0246***
	(0.006)	(0.007)	(0.008)
Year=2010	0.0147***	0.0440***	0.0019
	(0.005)	(0.008)	(0.007)
Year=2011	0.0113**	0.0350***	0.0009
	(0.006)	(0.007)	(0.007)
Sample size	271,478	110,688	160,790

Notes: Column (1) presents estimated coefficients for year dummies from full sample. Results in column (2) and (3) are from subsamples for children of college and non-college mothers, respectively. Standard errors are reported in brackets. In all regressions, we include for child's gender firstborn status, age group, race, indicator for those moved from a different state, and facility type, mother's marital status, education, and age group, uninsured share of children, immigration share of residence, estimated population, and state fixed effects.

Figure 2.2 Differences in MMR Non-Uptake Rate by Education



Notes: The top graph plots estimated coefficients of year dummies in column (2) and (3) of Table 2.4. Solid line is for college sample and dash line is for non-college sample. The bottom graph depicts the difference between estimates from college and non-college samples.

Table 2.5 Impact of Biased Beliefs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
college x ln(year)	0.0137***	0.0092**	0.0134***	0.0093**	0.0135***	0.0096**	0.0134***	0.0091**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
college x post	0.0014		0.0009		0.0007		0.0011	
	(0.005)		(0.005)		(0.005)		(0.005)	
college x 1{year=2004}		0.0056		0.0054		0.0054		0.0054
		(0.007)		(0.007)		(0.007)		(0.007)
college x 1{year=2005}		-0.0020		-0.0023		-0.0022		-0.0022
		(0.008)		(0.008)		(0.008)		(0.008)
college x 1{year=2006}		-0.0054		-0.0057		-0.0063		-0.0056
		(0.008)		(0.008)		(0.008)		(0.008)
college x 1{year=2007}		0.0102		0.0089		0.0085		0.0097
		(0.008)		(0.008)		(0.008)		(0.008)
college x 1{year=2008}		0.0064		0.0046		0.0039		0.0058
		(0.008)		(0.008)		(0.008)		(0.008)
college x 1{year=2009}		-0.0010		-0.0022		-0.0027		-0.0016
		(0.010)		(0.010)		(0.010)		(0.010)
college x 1{year=2010}		0.0257***		0.0242**		0.0240**		0.0250***
		(0.010)		(0.010)		(0.010)		(0.010)
college x 1{year=2011}		0.0156		0.0140		0.0131		0.0146
		(0.010)		(0.010)		(0.010)		(0.010)
Sample size	271,478	271,478	271,478	271,478	271,478	271,478	271,478	271,478
Area specific time trend	No		Region x year FE		Division x year FE		State x ln(year)	

Notes: The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. In the even columns, indicators for each year after 2004 are included instead of the post indicator. All controls in Table 4 are included. No area specific time trend is controlled in columns (1) and (2). We include region specific year effects in columns (3) and (4), division specific year effects in columns (5) and (6), and state specific log linear time trend in columns (7) and (8). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6 Mechanisms Underlying Biased Beliefs: Single Information Exposures

	(1)	(2)	(3)	(4)
College x ln(year)	0.0129***	0.0130***	0.0134***	0.0132***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0014	0.0008	0.0006	0.0011
	(0.005)	(0.005)	(0.005)	(0.005)
College x ln(year) x autism	0.0001	0.0004	0.0005	0.0010
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x reported cases	-0.0021***	-0.0018**	-0.0018**	-0.0020**
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x news counts	0.0008	0.0004	0.0003	0.0009
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news for vaccine	0.0046	0.0044	0.0039	0.0046
	(0.003)	(0.003)	(0.003)	(0.003)
College x ln(year) x news w/ science	-0.0034	-0.0026	-0.0023	-0.0032
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ authority	-0.0045**	-0.0042**	-0.0039**	-0.0040**
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ parents	0.0023	0.0021	0.0021	0.0023
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x autism search	0.0009	0.0001	-0.0001	-0.0003
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x outbreaks search	-0.0035***	-0.0033***	-0.0034***	-0.0038***
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x "vaccine autism" search	0.0036***	0.0030***	0.0030***	0.0030***
	(0.001)	(0.001)	(0.001)	(0.001)
Sample size	271,478	271,478	271,478	271,478
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Coefficients are reported for triple interaction terms using all the information exposures presented in Table 3. We use z-scores for each type of information exposure. All controls in Table 4 are included. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.7 Mechanisms for Biased Beliefs: Information Sources

	(1)	(2)	(3)	(4)
College x ln(year)	0.0117***	0.0118***	0.0121***	0.0119***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0014	0.0009	0.0007	0.0011
	(0.005)	(0.005)	(0.005)	(0.005)
Disease Prevalence rate				
College x ln(year) x harmful signal	-0.0029	0.0102	0.01733	0.0742
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x safe signal	-0.0022***	-0.0020**	-0.0020**	-0.0021**
	(0.001)	(0.001)	(0.001)	(0.001)
News coverage				
College x ln(year) x harmful signal	0.0020	0.0015	0.0014	0.0019
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x safe signal	-0.0014	-0.0007	-0.0008	-0.0006
	(0.002)	(0.002)	(0.002)	(0.002)
Online searches				
College x ln(year) x harmful signal	0.0057***	0.0046***	0.0045***	0.0040***
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x safe signal	-0.0018*	-0.0016	-0.0017	-0.0020*
	(0.001)	(0.001)	(0.001)	(0.001)
Sample size	271,478	271,478	271,478	271,478
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Coefficient of the triple interaction term using harmful signal of disease prevalence rate is multiplied by 100 for reporting purpose. For disease prevalence rate, the harmful signal is the z-score of autism prevalence rate; and the safe signal is the z-score of reported total cases of measles, mumps, and rubella. For news coverage, the harmful signal is the average z-scores of news count and percentage of news with parents' opinions; and the safe signal is the average z-scores of percentage of news encouraging immunization, with words from authorities, and scientific proofs. For web searches, the harmful signal is the z-score of search index for measles, mumps, and rubella; and the safe signal is the average of z-scores of search index for autism and "vaccine and autism" topics. All controls in Table 4 are included. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.8 Mechanisms for Biased Beliefs: Information Attitudes

	(1)	(2)	(3)	(4)
College x ln(year)	0.0111***	0.0112***	0.0115***	0.0112***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0014	0.0008	0.0007	0.0011
	(0.005)	(0.005)	(0.005)	(0.005)
College x ln(year) x composite harmful signal	0.0071***	0.0059***	0.0058***	0.0065***
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x composite safe signal	-0.0026	-0.0019	-0.0022	-0.0027
	(0.002)	(0.002)	(0.002)	(0.002)
Sample size	271,478	271,478	271,478	271,478
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Harmful composite is the average z-scores of autism prevalence rate, news count, percentage of news with parents' opinions, and search index for autism, and "autism and vaccine" topic. Safe composite is the average z-scores of reported total cases for measles, mumps, and rubella, percentage of news encouraging immunization, with words from authority, and scientific proofs, and search index for measles, mumps, and rubella. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.9 Impact of Biased Beliefs--Children not Moved from a Different State

	(1)	(2)	(3)	(4)
College x ln(year)	0.0122***	0.0118***	0.0120***	0.0116***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0037	0.0030	0.0029	0.0035
	(0.006)	(0.005)	(0.005)	(0.006)
Sample size	249,358	249,358	249,358	249,358
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The sample is restricted to children not moved from a different state. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. All controls in Table 4 are included. No area specific time trend is controlled in columns (1) and (2). We include region specific year effects in columns (3) and (4), division specific year effects in columns (5) and (6), and state specific log linear time trend in columns (7) and (8). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.10 Mechanism for Biased Beliefs: Single Information Exposures--Children not Moved from a Different State

	(1)	(2)	(3)	(4)
College x ln(year)	0.0116***	0.0118***	0.0121***	0.0116***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0036	0.0029	0.0028	0.0034
	(0.006)	(0.005)	(0.005)	(0.006)
College x ln(year) x autism	-0.0002	0.0001	0.0002	0.0006
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x reported cases	-0.0017**	-0.0015*	-0.0015*	-0.0017**
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x news counts	0.0009	0.0005	0.0003	0.0008
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news for vaccine	0.0048	0.0047	0.0044	0.0053
	(0.003)	(0.003)	(0.003)	(0.003)
College x ln(year) x news w/ science	-0.0038*	-0.0031	-0.0028	-0.0037
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ authority	-0.0043**	-0.0040**	-0.0038**	-0.0038*
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ parents	0.0028*	0.0026	0.0026	0.0027
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x autism search	0.0013	0.0004	0.0002	-0.0001
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x outbreaks search	-0.0036***	-0.0035***	-0.0035***	-0.0038***
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x "vaccine autism" search	0.0031***	0.0026***	0.0026***	0.0024**
	(0.001)	(0.001)	(0.001)	(0.001)
Sample size	249,358	249,358	249,358	249,358
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The sample is restricted to children not moved from a different state. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Coefficients are reported for triple interaction terms using all the information exposures presented in Table 3. We use z-scores for each type of information exposure. All controls in Table 4 are included. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.11 Mechanisms for Biased Beliefs: Information Sources--Children not Moved from a Different State

	(1)	(2)	(3)	(4)
College x ln(year)	0.0107***	0.0108***	0.0111***	0.0106***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0037	0.0030	0.0028	0.0035
	(0.006)	(0.005)	(0.005)	(0.006)
Disease Prevalence rate				
College x ln(year) x harmful signal	-0.0003	-0.0002	-0.0001	0.0004
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x safe signal	-0.0018**	-0.0016**	-0.0016**	-0.0018**
	(0.001)	(0.001)	(0.001)	(0.001)
News coverage				
College x ln(year) x harmful signal	0.0027*	0.0023	0.0021	0.0024
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x safe signal	-0.0017	-0.0010	-0.0010	-0.0004
	(0.002)	(0.002)	(0.002)	(0.002)
Online searches				
College x ln(year) x harmful signal	0.0053***	0.0041***	0.0040***	0.0035**
	(0.001)	(0.001)	(0.001)	(0.002)
College x ln(year) x safe signal	-0.0024**	-0.0021**	-0.0022**	-0.0025**
	(0.001)	(0.001)	(0.001)	(0.001)
Sample size	249,358	249,358	249,358	249,358
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The sample is restricted to children not moved from a different state. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. For disease prevalence rate, the harmful signal is the z-score of autism prevalence rate; and the safe signal is the z-score of reported total cases of measles, mumps, and rubella. For news coverage, the harmful signal is the average z-scores of news count and percentage of news with parents' opinions; and the safe signal is the average z-scores of percentage of news encouraging immunization, with words from authorities, and scientific proofs. For web searches, the harmful signal is the z-score of search index for measles, mumps, and rubella; and the safe signal is the average of z-scores of search index for autism and "vaccine and autism" topics. All controls in Table 4 are included. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.12 Mechanisms for Biased Beliefs: Information Attitudes--Children not Moved from a Different State

	(1)	(2)	(3)	(4)
College x ln(year)	0.0095**	0.0097***	0.0100***	0.0095***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0037	0.0030	0.0028	0.0035
	(0.006)	(0.005)	(0.005)	(0.006)
College x ln(year) x composite harmful signal	0.0061***	0.0042**	0.0040**	0.0042**
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x composite safe signal	-0.0018	-0.0007	-0.0008	-0.0007
	(0.002)	(0.002)	(0.002)	(0.002)
Sample size	249,358	249,358	249,358	249,358
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The Sample is restricted to children not moved from a different state. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Harmful composite is the average z-scores of autism prevalence rate, news count, percentage of news with parents' opinions, and search index for autism, and "autism and vaccine" topic. Safe composite is the average z-scores of reported total cases for measles, mumps, and rubella, percentage of news encouraging immunization, with words from authority, and scientific proofs, and search index for measles, mumps, and rubella. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

3. Hospital Report Cards and Hospital Attractiveness to Commercial HMO Insurers

3.1. Introduction

Since 1990s, government has been devoted numerous resources to publish hospital report card with quality ratings for individuals to identify high quality providers as well as to promote efficiency on the provider side. Around 2007, price transparency action has been called for as a tool to bring medical cost under control by encouraging individuals to choose medical providers with cost-consciousness. Currently, more than half of the states report at least somewhat price information either alone or combined with existing quality report card.

One major concern facing the price transparency action is what price to publish. There are three commonly reported prices, which differ in degrees of transparency and criticism they are facing (Sinaiko and Rosenthal 2011). The most commonly reported type is average charge by hospital. Charge is the list price of hospital stay, which is a poor proxy for the transaction price actually paid to the hospital. Besides, individual may consider charge as an indicator for quality because higher quality might be more costly. Then, whether the intention of price transparency will be achieved depends on how the majority individuals perceive the charge information, which may affect the relative bargain position between hospitals and insurers, providing another mechanism for the published price information to play a role.

This paper we use data from Pennsylvania to study whether published hospital report card information, including both quality ratings and average charge, for Coronary Artery Bypass Graft (CABG) surgery is correlated with hospital attractiveness to commercial HMO insurer. Pennsylvania CABG surgery market provides a good setting for us because average charge information along with quality ratings is reported annually since 1998, and there is no other type

of price information available in the state. The impact of report card information on the behavior of individual (Wang et al. 2011) and health care providers (Dranove et. al. 2003; Kolstad 2013) has been widely studied, but with a focus on quality ratings only. Moreover, researchers have not directly studied the effect of report card information on insurers, making our study a supplement to the existing literature.

We consider our research question under the following framework. Individuals choose insurance plan that maximize their utility, which depends positively on the expected utility they are able to gain from the hospital network of the insurer, given the insurance premium they have to pay. For individual insured by HMO plans, he may visit only the hospitals in that plan's network. On the other hand, insurers maximize their profit by having more enrollees, given their reimbursement rate. Therefore, which hospital to include in the network indirectly affects the attractiveness of the plan to individuals, and thus the profitability of the plan. A hospital will be more attractive to the insurer if individual shows strong preference toward the hospital, it might not be so if the hospital is substitutable by the existing hospitals in the insurer's network. We assume that such incentive is stronger for commercial plans.

Based on the above framework, we construct a proxy for hospital relative attractiveness to insurer by aggregating individual willingness-to-pay for a commercial HMO insurer if a particular hospital is included in its network using Pennsylvania Health Care Cost Containment Council (PHC4) inpatient discharge data. We exams whether hospital report card information is associated with the relative bargaining position between hospitals and insurers. We further provide a discussion on the likelihood that changes in bargaining relationship measured by hospital attractiveness will eventually result in adjustment of the network in a given plan.

The paper outlines as follows. In section 3.2, we provide a literature review. In section 3.3, we describe the data source. In section 3.4, we lay out empirical application. In section 3.5, we provide results and discussion. In section 3.6, we perform robustness check. In section 3.7, we present conclusions.

3.2. Literature Review

There is a large body of literature examining the impact of hospital report card. Studies that rely on observed consumer behavior find little effect of quality reporting at hospital level (Wang et al. 2011) and more evidence, though effects are relatively small, on aggregate market share (Dranove and Stefka 2008). Another line of research focuses on the behavior of health care provider by investigating selection against sicker patients (Dranove et al. 2003). Both of the lines of study focus on published quality ratings only. Our analysis complete the existing study by incorporating reported charge information into discussion and consider the effect of report card information on the bargaining position between hospitals and insurers, which has not been studied.

Studies investigating the bargaining between hospitals and insurers focus on the impact of market competition on the negotiated price paid to hospitals. Town and Vistnes (2001) investigate the determinants of actual negotiated prices paid to hospitals by two major HMOs in Los Angeles area from 1990-1993 and find that hospital bargain power decreases when HMO can readily turn to alternative networks that exclude the hospital. Using detailed California claims data, Ho and Lee (2013) conclude that increasing insurer competition lower prices on average but the most attractive hospitals can leverage increased competition to negotiate higher rates. In our study, we do not observe insurance characteristics. Instead, we measure the bargain

relationship solely based on aggregated change in expected utility of individual if a hospital is included in the network of a particular insurer.

3.3. Data Source

This paper pulls together information from several datasets. The first is the Pennsylvania Inpatient Hospital Discharge Data collected by PHC4. This dataset provides rich patient-level information on patient's demographics, zip code of residence, health insurer, admission type, the quarter of admission, diagnosis and procedure details, and a four-digit unique hospital identifier. Hospital characteristics were taken from the second data source, American Hospital Association's Annual Survey of Hospitals. We consider only CABG records and commercial health plan enrollees.

The third dataset we use in our study is the Pennsylvania's Guide to Coronary Artery Bypass Graft Surgery, which is referred as CABG report cards. The report card publication date and each patient's admission year and quarter allow us to identify inpatient visits during the same report card episodes. Each report card publish one of the following ratings--"higher than expected", "as expected", and "lower than expected"--for in-hospital mortality, 30-day mortality, 7-day readmission rate, and 30-day readmission rate for CABG hospitals. It also reports average post-surgical length of stay and total charge per patient. Consistent with previous studies, we use in-hospital mortality as a major quality measure. We define a hospital is of superior (inferior) if in-hospital mortality was rated as "lower than expected" ("higher than expected") in the most recent report card. In order to incorporate rating information on both mortality and readmission rate, we create a set of composite scores as an alternative measure for hospital quality: we define a hospital as of inferior if it received at least one "higher than expected" rating, irrespective of its ratings in other categories; and we define a hospital as of superior if it received at least one

"lower than expected" and no "higher than expected" ratings.[51] Accordingly, we re-define charge into categorical variables. We identify a hospital as of high charge if its average charge is in the upper quartile, and as of low charge if its average charge falls into the lower quartile in the most recent report card.[52] Table 3.1 lists the number of hospitals by ratings and average charge across four report card episodes, as well as the sample years of PHC4 inpatient discharge data that are matched to each report card. During the study period, the number of CABG hospitals is quite constant. Our composite single scores provide more variations in hospital quality than in-hospital mortality ratings only.

3.4. Empirical Application

3.4.1. Hospital Demand

We use a discrete choice model that allows for observed differences across individuals to estimate demand for hospitals during each CABG report card period. With some probability, individual i needs CABG surgery. His utility from choosing hospital j is given by

$$u_{ij} = \emptyset_j + h_j\alpha + h_j x_i \beta + \varepsilon_{ij}, \tag{3.1}$$

where $\emptyset_j$, h_j are unobserved and observed hospital characteristics, respectively; x_i is observed individual characteristics, and ε_{ij} is an idiosyncratic error term assumed to be i.i.d. Type I extreme value. Hospital characteristics include number of doctors per bed, and total bed size. Individual characteristics include gender, an indicator of age above 65, emergency status at admission, severity status (Charlson score greater than zero). In addition, we also include

[51] During our study period, two mortality ratings are the same for all hospitals, as are the two ratings for readmission rate. However, there are variations in ratings across the two general categories. For example, some hospitals received "lower than expected" for mortality but "higher than expected" for readmission, or verse visa.

[52] We also create another set of composite scores for quality as a robustness check by incorporating reported information on length of stay. We assign "higher than expected" for hospitals with length of stay in the upper quartile, and "lower than expected" for hospitals with length of stay in the bottom quartile in the most recent report card. We then define superior and inferior for each hospital using the same definition as described above.

individual travel distance to the hospital, distance squared, and interactions between these and individual characteristics.[53] Equation (3.1) is estimated using maximum likelihood techniques and PHC4 inpatient data for each report card period.

Since the choice sets of commercial HMO enrollees are unobserved, we consider the choices made by commercial FFS and PPO enrollees, whose choice set is unrestricted. We assume that commercial FFS/PPO enrollees have the same preferences over hospitals as commercial HMO enrollees, conditional on their observed characteristics. That is, we use the estimated coefficients from the hospital demand equation for commercial FFS/PPO enrollees to predict commercial HMO enrollees' favor for hospitals.

Table 3.2 reports conditional logit estimates of hospital demand for commercial FFS/POS enrollees during each report card period. Each individual in the regression sample has the same choice set--all CABG hospitals operating in that report card period. In each column, the distance coefficient is negative and highly significant. The results suggest that higher-demanding patients, such as relatively severe or emergent, are less distance sensitive.

3.4.2. Willingness-To-Pay

Following previous literature (Ho and Lee 2013), we use changes in enrollee's willingness-to-pay (WTP) when a hospital is added to a particular plan's network as a proxy for the relative attractiveness of the hospital to that insurer. It is one of the important determinants for the bargaining outcome between hospital and insurer.

We first use the estimated coefficients from the hospital demand equation to predict the utility of HMO enrollees provided by each hospital network of HMO plans. We focus on most common

[53] In order to estimate the distance between an individual and the various hospitals in his choice set, we obtained longitude and latitude of each hospital based on its street address and those of each patient based on his home zip code using user-written "geocode3" command in Stata. We then calculate travel distance between the two points in miles using user-written "geodist" command in Stata.

health plans listed in the Appendix C of Pennsylvania Uniform Claims and Billing Form Reporting Manual. Consistent with Ho and Pakes (2014), we infer the hospital network of each commercial HMO insurer using PHC4 inpatient data by assuming that a hospital is in the network if at least three patients are admitted from the particular insurer. Table 3.3 displays the commercial HMO plans studied in this paper along with the number of hospitals in its network during each report card episode. Then, individual i's expected utility from the hospital network offered by plan k when he needs CABG surgery can be written as

$$EU_{ik} = \log(\sum_{j\epsilon P_k} \exp(\widehat{\emptyset}_j + h_j\hat{\alpha} + h_j x_i\hat{\beta})), \tag{3.2}$$

where P_k is the set of hospitals offered to enrollees by plan k for CABG surgery. The change in expected utility from having hospital $j\epsilon P_k$ in the network is then given by

$$\Delta EU_{ijk} = EU_{ik} - EU_{ik}(P_k \backslash jk). \tag{3.3}$$

Prior to enrolling in health plan k and before individual i is sick, individual i's expected utility from having hospital $j\epsilon P_k$ in the network is then given by

$$WTP_{ijk} = \rho_i \Delta EU_{ijk}, \tag{3.4}$$

where ρ_i is individual i's probability of admission to any hospital for CABG surgery conditional on single age, gender, and year in Pennsylvania.[54] We then aggregate this measure over all HMO enrollees during a given report card period to obtain a measure of hospital j's attractiveness to plan k in our final analysis, which is given by

$$WTP_{jk} = \sum_i \rho_i \Delta EU_{ijk}. \tag{3.5}$$

As noted by Ho and Lee (2013), this measure captures not only hospital quality but also substitutability in the plan network. The relative differences across hospitals' WTP in a

[54] The probability is estimated using PHC4 inpatient data and population by single age, gender and year in Pennsylvania from http://seer.cancer.gov/popdata.

particular plan depend on characteristics of both the hospital itself and the other hospitals in the same network.

3.4.3. Main Regression Equation

The research question we interested in is whether public report cards information will affect the relative attractiveness of a hospital to an insurer in the existing network. In our final regression we consider hospital j in the network of plan k during report card episode r

$$WTP_{jkr} = \beta_1 ratings_{jr} + \beta_2 charge_{jr} + \tau_r + \gamma h_j + \varphi_k + v_{ijk}, \quad (3.6)$$

where the outcome variable is the hospital j's relative attractiveness to insurer k in the hospital report card period r; $ratings_{jr}$ and $charge_{jr}$ are report card information on quality and patient charge; τ_r is report card episode fixed effects; h_j is observed hospital characteristics, including non-for-profit status, indicator of teaching council member, doctors per bed, and bed size; and φ_k is plan fixed effects.[55] Coefficients of interest are β_1 and β_2.The impact of report card information on hospital relative attractiveness to insurer is identified by variations in reported ratings and charge across hospitals in the same network.

3.5. Results and Discussion

3.5.1. Main Results

Table 4 provides statistics summary for the full sample and each subsample within a given report card episode. Overall, hospital attractiveness to insurer and the size of insurer's network decrease over time. In contrast, reported charge increases across the four episodes and reported ratings for both in-hospital mortality and composite single score fluctuate.

To characterize the impact of report card information on hospital relative attractiveness to insurer, we estimate equation (6). Table 3.5 reports coefficients for reported ratings and charge,

[55] We use the most recent value for hospital characteristics during each report card episode. For example, for report card episode 2009/3 -2010/4, we obtain AHA hospital information in 2010.

and interactions between these two. In column (1)-(4), ratings are defined based on in-hospital mortality: superior (inferior) is a dummy equal to one if in-hospital mortality is lower (higher) than expected in the report card episode; in column (5)-(6), ratings are defined based on composite single score. Other controls are included step by step. Column (1) and (5) include report card fixed effects; Column (2) and (5) include hospital characteristics, including non-for-profit status, indicator for member of teaching council, doctors per bed, and bed size; Column (3), (4), (7), and (8) further include plan fixed effects.

We find that coefficients for charge variables are consistent with our expectations in signs and consistently significant. In contrast, all coefficients for report card ratings are insignificant, although the signs are generally as predicted. Compared to the sample mean, high charge hospitals are 53.4% less attractive to insurer, and low charge hospitals are 76.1% more attractive to insurer given the network of the plan, which suggests a quite substantial impact of reported charge on hospital relative attractiveness. Among variables for hospital characteristics, only hospital bed size remains positive and significant at 5 percent level across all specifications. It suggests that big hospitals enjoy more provider leverage in contract negotiation as they may generate more patient volume for a given insurer (Berenson et al., 2012).

In column (4) and (8), we add interaction terms between reported ratings and charge with the most inclusive specification. The estimate for high charge interacted with in-hospital mortality lower than expected is not available, because there is no observation in this category in our final sample. Upon the inclusion of these interactions, main effects of rating and charge variables do not change materially, and all the interaction terms are insignificant. The result suggests that there is no interacted impact of ratings and charge on hospital relative attractiveness.

We expect that plans with a small hospital network are affected by the reported charge more because patient volumes are guaranteed by more hospitals in large networks making relative importance of one single hospital to be less. To test our hypothesis, we define a plan has small network if number of hospitals in its network is in the bottom quartile of the whole sample. And we include interaction terms between the small network indicator and report card information in regressions. Results are reported in Table 3.6. In column (1) and (2), superior and inferior ratings are defined based on in-hospital mortality; in contrast, in column (3) and (4), they are defined based on composite single score. Consistent with above findings, both plan types show distastes for high charge hospitals at 5-percent level. In contrast, only small plan favors low charge hospitals. Compared to the mean, a low charge hospital in the most recent report card episode is around 80% more attractive to the plan. In contrast, attractiveness only increases by 30% for low charge hospitals in non-small network and the estimate is not statistically different from zero.[56] The result also suggests that superior hospitals are more attractive in small sized network compared to non-small ones, although the findings are not consistent when we consider in-hospital mortality only.

3.5.2. Discussion

The above findings suggest that reported charge will affect hospital attractiveness to insurer in a given plan. Overall, insurers prefer low charge hospitals and dislike high charge hospitals. Impact from report card ratings is weak and only exists in small network for superior hospitals, if any. However, whether the impacts on hospital relative attractiveness to insurers will eventually result in inclusion or exclusion of hospitals for a given plan still remains uncertain. Since we do not observe actual change of plan's network over year, we have to infer it based on our implied

[56] The means for willingness-to-pay is 1.189 for small sized network and 0.371 for non-small sized network. The impact of low charge on hospital attractiveness in percentage term is around 80% ((0.145+0.185)/1.189) in small sized network and about 30% (0.145/0.371) in non-small sized network.

network, which may introduce measurement errors. We define a hospital is added into a given plan if the hospital does not exist in the plan's network during last report card episode. Accordingly, we define a hospital is dropped from a given plan if the hospital exist in the plan's network during the following report card episode, but no longer does in the current report card episode. Then, we consider, for hospital j in plan k during report card episode r

$$inclusion_{jkr} = \theta WTP_{jkr} + \tau_r + \gamma h_j + \emptyset_j + \varphi_k + \mu_{jkr}, \quad (3.7)$$

$$exclusion_{jkr} = \theta WTP_{jkr} + \tau_r + \gamma h_j + \emptyset_j + \varphi_k + \mu'_{jkr}, \quad (3.8)$$

where $inclusion_{jkr}$ is a dummy equal to one if the hospital is newly included in the network during the current report card period, $r-1$; $exclusion_{jkr}$ is a dummy equal to one if the hospital is dropped from the network in the next report card period, $r+1$; consistent with above, WTP is willingness-to-pay at hospital-plan level, τ_r is report card episode fixed effect, h_j is hospital observed characteristics, $\emptyset_j$ is hospital fixed effects, and φ_k is plan fixed effects. The coefficient of interest is θ, which is expected to be positive in equation (7) and negative in equation (8). Because we predict that a given hospital is more likely to be added into a plan that values it more and more likely to be dropped from a plan with distaste for it.

Table 3.7 reports estimates from equation (7) in panel A and those from equation (8) in panel B. Other control variables are included step by step from column (1)-(3). We find that as relative attractiveness to insurer increases, a given hospital is more likely to have more HMO contracts with commercial insurer. The impact is significant as expected. In the most inclusive specification, the hospital is 4.22 percentage points more likely to be included in a plan's network in response to one unit increase in willingness-to-pay, which corresponds to an impact of 27% in percentage term.[57] However, contradicting to our expectations, willingness-to-pay fails

[57] Average probability of hospital inclusion in network in the sample is 0.156.

to explain the exclusion of hospitals from plans because the estimates in panel B are negligible and insignificant across all specifications. One possible explanation is that employer are resistant to choice-limiting networks with few providers. Therefore, plans may lack an important bargaining chip, without a credible threat of excluding a provider from their network (Berenson et al. 2012). Based on our estimation, hospital with reported low charge are 20.5% (=76.4% x 27%) more likely to be added into an HMO plan's network by changing the relative attractiveness to its insurer.

3.6. Robustness

We perform two types of robustness check. First, we replicate our main results using superior and inferior ratings that include post-surgical length of stay. Results are reported in Table 3.8. Our findings are generally robust to this change, although estimates lose significance when full interaction terms between reported ratings and charge are included.

Second, we check the sensitivity of our results by changing the definition of the implied network. Column (1) of Table 3.9 reports estimates from the sample that assume a hospital is in the network of a particular insurer if at least 1 patient is admitted from that insurer, which is a broader definition but it may also wrongly counts patients who go out-of-network into a plan's network; and column (2) displays estimates form the sample that assume a hospital is in the network of a particular insurer if at least 5 patients are admitted from that insurer, which is a narrower definition but it may be too selective given the small patient volume of CABG surgery from commercial enrollees in each report card episode. Panel A reports estimated coefficients of equation (6); Panel B displays estimated coefficients of equation (7); and Panel C provides estimated coefficients of equation (8). We find that results in panel A are not sensitive to the change of network definition. Although the impact of high charge on hospital relative

attractiveness is not significant, impact of low charge are negative and significant. More hospital attractiveness to insurer lead to higher probability of inclusion in the plan, however, the estimates are not significant if we use narrower definition of plan network.

3.7. Conclusion

This paper investigates whether published report card information will affect hospital relative attractiveness to commercial HMO insurer for CABG surgery in Pennsylvania over four report card episodes from 2006-2010. We consider impact of both reported hospital ratings and charge. We use change in expected willingness-to-pay for the plan if a hospital is dropped from the network of a particular insurer as a proxy for hospital relative bargaining position against insurer, given the characteristics of the plan. We hypothesis that report card information exerts an impact on the bargaining position between hospital and insurer as it affect the patient flow a provider will generate to the insurer.

We first estimate hospital demand among commercial FFS/PPO enrollees during each report card episode. We then use the estimates to imply the tastes of commercial HMO enrollees. We calculate change in consumer expected utility of holding a particular plan if a hospital is excluded from its network using implied network from PHC4 inpatient discharge data. A hospital is assumed in the network of a particular insurer if at least three patients are admitted from the insurer. We finally aggregate over patients to obtain the outcome variable of interest, willingness-to-pay at hospital-plan level. Our results suggest that reported charge is strongly associated with hospital relative attractiveness. Given the insurance plan, low charge hospitals are more attractive and thus less likely to be substitute by the other alternatives. Based on our calculation, hospitals with reported low charge are 20.5% more likely to be added into the network of an HMO plan due to changes in the relative attractiveness to its insurer.

Our study has several limitations. First, we do not observe the actual network of commercial HMO plans. The implied network may introduce measurement errors. Second, characteristics of the insurer are not available to us, such as premiums, number of enrollees, and rated performance of the plan, all of which may also affect the attractiveness of the plan to patient, and thus affect negotiation outcomes between hospital and insurer. Third, we are not able to adequately explain the reason for plans to drop a hospital from their network. These are possible directions to improve in the future.

Table 3.1 Hospital Report Card Ratings and Average Charge

Publication year/quarter	2009/3	2008/3	2007/2	2006/1
Total number of hospitals	60	59	60	60
Number of hospitals by in-hospital mortality ratings				
Superior	1	5	6	5
Inferior	4	3	2	4
Number of hospitals by composite single scores				
Superior	2	5	8	9
Inferior	11	12	9	11
Hospital average charge in thousands				
Mean	126.018	117.736	113.522	107.270
Standard deviation	(76.871)	(73.234)	(71.856)	(67.523)
Data collection year	2007	2006	2005	2004
Report card period matched to	2009/4-2010/3	2008/4-2009/3	2007/3-2008/3	2006/2-2007/2

Note: The data source is Pennsylvania's Guide to Coronary Artery Bypass Graft Surgery. CABG hospitals with quality ratings in each report card episode are included. For in-hospital mortality ratings, superior (inferior) is defined as a hospital received "lower than expected" ("higher than expected") rating in the most recent report card. Composite scores combine report card grades on in-hospital mortality, 30-day mortality, 7-day readmission, and 30-day readmission. A superior rating means a hospital performed above expectation in at least one of the categories, and below expectation in none. An inferior rating means a hospital performed below expectation in at least one category. Hospital charges are in thousands and unadjusted for inflation.

Table 3.2 Hospital Demand Estimates from A Conditional Logit Model

	(1)	(2)	(3)	(4)
Report card episode	2009/4-2010/3	2008/4-2009/3	2007/3-2008/3	2006/2-2007/2
Doctors per bed	0.4745	-0.5499	-3.2886***	-0.5513
	(1.036)	(1.701)	(1.207)	(0.861)
Bed size	0.1338	-0.1431	-0.1881	0.261
	(0.248)	(0.329)	(0.17)	(0.233)
Distance	-0.1696***	-0.1538***	-0.1517***	-0.1398***
	(0.010)	(0.009)	(0.008)	(0.008)
Distance squared	4.8594***	1.2637***	1.7436***	-0.0249
	(0.446)	(0.485)	(0.486)	(0.235)
Severity x doctors per bed	0.0747	0.3331	0.7935***	0.3675
	(0.211)	(0.233)	(0.215)	(0.227)
Severity x bed size	-0.04	-0.0289	-0.0438*	0.0114
	(0.030)	(0.025)	(0.023)	(0.023)
Severity x distance	0.0213***	0.0176***	0.0200***	-0.0083*
	(0.007)	(0.006)	(0.005)	(0.005)
Severity x distance squared	-1.3082***	-0.2361	-0.8246***	0.321
	(0.374)	(0.179)	(0.162)	(0.223)
1{age>=65} x doctors per bed	0.2702	0.2354	-0.1954	0.1302
	(0.269)	(0.277)	(0.270)	(0.262)
1{age>=65} x bed size	-0.0469	-0.0068	-0.0039	-0.0176
	(0.042)	(0.035)	(0.032)	(0.031)
1{age>=65} x distance	0.0068	-0.004	0.0124**	0.0067
	(0.005)	(0.008)	(0.006)	(0.005)
1{age>=65} x distance squared	0.0247	-0.8430**	-0.6598*	0.2122***
	(0.069)	(0.333)	(0.347)	(0.053)
Male x doctors per bed	-0.0537	0.1205	0.1513	0.3108
	(0.250)	(0.261)	(0.241)	(0.269)
Male x bed size	0.0691*	-0.0186	-0.0451*	0.005
	(0.039)	(0.030)	(0.025)	(0.028)
Male x distance	0.0138*	0.0157**	0.0023	0.0258***
	(0.008)	(0.007)	(0.007)	(0.007)
Male x distance squared	-0.8702**	-0.7649*	-0.6658	0.1926
	(0.396)	(0.451)	(0.481)	(0.232)
Emerge x doctors per bed	-0.0024	-0.0828	0.0896	0.0922
	(0.193)	(0.204)	(0.188)	(0.197)
Emerge x bed size	-0.032	-0.0429*	-0.0881***	-0.0830***
	(0.028)	(0.023)	(0.021)	(0.021)
Emerge x distance	0.0249***	0.0049	0.0002	0.0168***
	(0.006)	(0.005)	(0.005)	(0.005)

Emerge x distance squared	-2.4400***	0.5115***	1.5566***	0.4893**
	(0.338)	(0.079)	(0.198)	(0.242)
Sample size	114,780	120,000	156,529	161,711
Hospital fixed effects	Yes	Yes	Yes	Yes

Notes: Standard errors are reported in parentheses. Each column is a separate regression using samples from PHC4 inpatient discharge during the corresponding report card episode. Regression samples only include commercial HMO/PPO enrollees. Distance is the miles between the zip code of patient's residence and hospital in his choice set. Severity is a dummy equal to one if patient's Charlson index greater than zero. And emerge is a dummy equal to one if patient's admission type is emergent and urgent. The specification includes hospital fixed effects. *significant at 10%; ** significant at 5%; ***significant at 1%.

Table 3.3 Implied Network of Commercial HMO Insurer in Pennsylvania

Plan	NAIC	Number of hospitals in the network during report card episode (year/quarter) 2009/4-2010/3	2008/4-2009/3	2007/3-2008/3	2006/2-2007/2
Aetna Health, Inc	95109	21	27	36	30
AmeriHealth	95044	1	1	2	2
CIGNA Healthcare of PA	95121	6	3	7	1
Geisinger Health Plan	95923	7	8	9	10
HealthAmerica (Central and Pittsburgh)	95060	10	10	15	16
HMO of Northeastern Pennsylvania	96601	5	6	6	5
Horizon Healthcare PA	95359	0	0	1	1
Keystone Health Plan Central, Inc.	95199	5	9	10	10
Keystone Health Plan East, Inc.	95056	22	23	22	27
Keystone Health Plan West, Inc.	95048	12	9	12	10
Optimum Choice, Inc. of PA	95225	0	0	0	1
UPMC Health Plan, Inc.	95216	9	10	10	10

Notes: Only most common health insurer and HMO plans are included in the table. Sizes of hospital network reported are implied using PHC4 inpatient discharge data during each report card episode by assuming that a hospital is in the network if at least three patients are admitted form the particular insurer.

Table 3.4 Sample Statistics

		Report card episode (year/quarter)			
	Full sample	2009/3-2010/4	2008/3-2009/4	2007/3-2008/3	2006/2-2007/2
	(1)	(2)	(3)	(4)	(5)
WTP	0.573	0.398	0.620	0.602	0.647
	(1.352)	(0.811)	(1.440)	(1.388)	(1.579)
Superior rating based on					
In-hospital mortality	0.050	0.010	0.069	0.032	0.085
	(0.218)	(0.101)	(0.254)	(0.177)	(0.280)
Composite single score	0.118	0.041	0.127	0.129	0.161
	(0.323)	(0.199)	(0.335)	(0.337)	(0.369)
Poor rating based on					
In-hospital mortality	0.084	0.082	0.098	0.089	0.068
	(0.277)	(0.275)	(0.299)	(0.285)	(0.252)
Composite single score	0.158	0.163	0.216	0.121	0.144
	(0.366)	(0.372)	(0.413)	(0.327)	(0.353)
Charge	111.059	120.739	116.545	104.94	104.709
	(19.934)	(20.764)	(19.839)	(20.951)	(17.010)
Non-for-profit	0.943	0.949	0.931	0.952	0.941
	(0.231)	(0.221)	(0.254)	(0.215)	(0.237)
Doctors per bed	0.181	0.166	0.171	0.182	0.199
	(0.246)	(0.240)	(0.243)	(0.248)	(0.253)
Bed size	4.428	4.578	4.499	4.377	4.297
	(2.224)	(2.430)	(2.314)	(2.208)	(1.989)
Member of teaching council	0.396	0.357	0.412	0.395	0.415
	(0.490)	(0.482)	(0.495)	(0.491)	(0.495)
Number of hospitals in the network	16.941	14.143	16.216	18.5	18.254
	(9.300)	(6.867)	(8.746)	(10.687)	(9.457)
Sample size	442	98	102	124	118

Notes: Mean coefficients are reported for the full sample and subsamples for each report card episode. Standard deviations are reported in parentheses. Superior rating based on in-hospital mortality is the rating of lower than expected; Inferior rating based on in-hospital mortality is the rating of higher than expected; the composite single score combines report card ratings on in-hospital mortality, 30-day mortality, 7-day readmission, and 30-day readmission. Superior rating based on composite single score is defined as above expectation in at least one of the categories, and below expectation in none; Inferior rating based on composite single core is defined as below expectation in at least one category. Charge is in thousand dollars and unadjusted for inflation.

Table 3.5 Willingness-To-Pay and Report Card Information

	In-hospital mortality				Composite single score			
Outcome=WTP	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High charge	-0.369**	-0.418**	-0.306*	-0.344*	-0.370**	-0.426**	-0.321*	-0.401*
	(0.164)	(0.184)	(0.184)	(0.191)	(0.164)	(0.183)	(0.182)	(0.206)
Low charge	0.393***	0.399**	0.436***	0.397**	0.382**	0.386**	0.438***	0.321*
	(0.149)	(0.156)	(0.154)	(0.161)	(0.149)	(0.158)	(0.155)	(0.180)
Superior	0.409	0.383	0.333	0.225	0.128	0.0877	0.00559	-0.177
	(0.295)	(0.295)	(0.273)	(0.332)	(0.201)	(0.206)	(0.191)	(0.264)
Inferior	-0.086	-0.0267	-0.0537	-0.219	-0.164	-0.138	-0.121	-0.297
	(0.232)	(0.233)	(0.216)	(0.330)	(0.177)	(0.178)	(0.164)	(0.249)
High charge x superior								0.195
								(0.925)
High charge x inferior				0.305				0.279
				(0.483)				(0.389)
Low charge x superior				0.330				0.415
				(0.584)				(0.396)
Low charge x inferior				0.269				0.370
				(0.705)				(0.436)
Sample size	443	443	443	443	443	443	443	443
Report card fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Hospital characteristics	no	yes	yes	yes	no	yes	yes	yes
Plan fixed effects	no	no	yes	yes	no	no	yes	yes

Notes: Standard errors are reported in parentheses. Each column is a separate regression. The dependent variable is hospital attractiveness to insurer at hospital-plan level. High charge is a dummy equal to one if charge reported in the most recent report card episode is in the top quartile; Low charge is a dummy equal to one if charge reported in the most recent report card episode is in the bottom quartile. In column (1)-(4), report card ratings are defined based on in-hospital mortality, that is, the hospital is considered as superior (inferior) if in-hospital mortality in a given report card episode is lower (higher) than expected. In column (5)-(6), report card ratings are defined based on composite single score, which combines report card ratings on in-hospital mortality, 30-day mortality, 7-day readmission, and 30-day readmission. Superior rating based on

composite single score is defined as above expectation in at least one of the categories, and below expectation in none; Inferior rating based on composite single core is defined as below expectation in at least one category. Other control variables are included step by step. Column (1) and (5) include report card fixed effects; Column (2) and (5) include hospital characteristics, including non-for-profit status, indicator for member of teaching council, doctors per bed, and bed size; Column (3), (4), (7), and (8) further include plan fixed effects. In our sample, there is no observation in the category of high charge and in-hospital mortality lower than expected. *significant at 10%; ** significant at 5%; ***significant at 1%.

Table 3.6 Willingness-To-Pay and the Size of Plan's Implied Network

	In-hospital mortality		Composite single score	
Outcome=WTP	(1)	(2)	(3)	(4)
High charge	-0.400**	-0.452**	-0.378**	-0.429**
	(0.167)	(0.185)	(0.166)	(0.184)
Low charge	0.102	0.076	0.145	0.111
	(0.168)	(0.177)	(0.171)	(0.181)
Superior	0.221	0.236	-0.127	-0.126
	(0.348)	(0.347)	(0.236)	(0.240)
Inferior	-0.137	-0.0616	-0.221	-0.19
	(0.260)	(0.261)	(0.193)	(0.193)
Small size x high charge	0.954	0.814	0.907	0.768
	(0.604)	(0.604)	(0.606)	(0.606)
Small size x low charge	0.929***	0.976***	0.815***	0.868***
	(0.260)	(0.261)	(0.270)	(0.271)
Small size x superior	0.590	0.489	0.878**	0.809**
	(0.603)	(0.601)	(0.397)	(0.396)
Small size x inferior	0.257	0.157	0.264	0.224
	(0.535)	(0.536)	(0.411)	(0.411)
Sample size	443	443	443	443
Report card fixed effects	yes	yes	yes	yes
Hospital characteristics	no	yes	no	yes

Notes: Standard errors are reported in parentheses. Each column is a separate regression. The dependent variable is hospital willingness-to-pay at hospital-plan level. High charge is defined as hospital charge in the top quartile in the report card episode; low charge is defined as hospital charge in the bottom quartile in the report card episode. In column (1) and (2), report card ratings are defined based on in-hospital mortality, that is, the hospital is considered as superior (inferior) if in-hospital mortality in a given report card episode is lower (higher) than expected. In column (3) and (4), report card ratings are defined based on composite single score, which combines report card ratings on in-hospital mortality, 30-day mortality, 7-day readmission, and 30-day readmission. Superior rating based on composite single score is defined as above expectation in at least one of the categories, and below expectation in none; Inferior rating based on composite single core is defined as below expectation in at least one category. Small size is a dummy equal to one if number of hospitals in the plan is in the bottom quartile of the whole sample. For the other controls, column (1) and (3) include report card fixed effects; column (2) and (4) further include hospital characteristics, including non-for-profit status, indicator for member of teaching council, doctors per bed, and bed size. *significant at 10%; ** significant at 5%; ***significant at 1%.

Table 3.7 Change of Plan's Implied Network and Willingness-To-Pay

	(1)	(2)	(3)
Panel A: outcome=probability of inclusion			
WTP	0.0312*	0.0725***	0.0422**
	(0.016)	(0.020)	(0.020)
Sample size	326	326	326
Panel B: outcome=probability of exclusion			
WTP	0.0037	0.0217	-0.0018
	(0.015)	(0.018)	(0.019)
Sample size	347	347	347
Report card episode fixed effects	yes	yes	yes
Hospital characteristics and hospital fixed effects	no	yes	yes
Plan fixed effects	no	no	yes

Notes: Standard errors are reported in the parentheses. Each column of each panel is a separate regression. In panel A, the dependent variable is a dummy equal to one if the hospital is newly included into a given network compared to last report card episode. In panel B, the dependent variable is a dummy equal to one if the hospital is dropped in a given network compared to the following report card episode. Estimated coefficients for the variable WTP are reported. Other control variables are included step by step. In column (1), report card episode fixed effects are included; in column (2), hospital fixed effects and hospital characteristics, including non-for-profit status, indicator for member of teaching council, doctors per bed, and bed size, are included. In column (3), plan fixed effects are further included. *significant at 10%; ** significant at 5%; ***significant at 1%.

Table 3.8 Robustness Check A

Outcome=WTP	(1)	(2)	(3)
High charge	-0.319*	-0.443	-0.401**
	(0.189)	(0.294)	(0.195)
Low charge	0.439***	0.329	0.072
	(0.155)	(0.234)	(0.181)
Superior	0.004	-0.143	0.040
	(0.153)	(0.207)	(0.185)
Inferior	-0.030	-0.068	-0.137
	(0.146)	(0.200)	(0.166)
High charge x superior		0.336	
		(0.789)	
High charge x inferior		0.122	
		(0.354)	
Low charge x superior		0.321	
		(0.332)	
Low charge x inferior		0.012	
		(0.358)	
Small size x high charge			0.774
			(0.649)
Small size x low charge			0.905***
			(0.291)
Small size x superior			0.212
			(0.283)
Small size x inferior			0.066
			(0.31)
Sample size	443	443	443
Report card episode fixed effects	yes	yes	yes
Hospital characteristics	yes	yes	yes
Plan fixed effects	yes	yes	no

Notes: Standard errors are reported in parentheses. Each column is a separate regression. The dependent variable is willingness-to-pay at hospital-plan level. High charge (low charge) is a dummy equal to one if hospital charge is in the top (bottom) quartile in the report card episode. Superior and inferior ratings are defined based on composite single score that includes post-surgical length of stay. Superior is a dummy equal to one if at least one of the categories is above expectation (in the top quartile for post-surgical length of stay), and below expectation (in the bottom quartile for post-surgical length of stay) in none; Inferior rating based on composite single core is defined as below expectation in at least one category. Small size is a dummy equal to one if number of hospitals in the plan is in the bottom quartile of the whole sample. *significant at 10%; ** significant at 5%; ***significant at 1%.

Table 3.9 Robustness Check B

	Define plan's hospital network if number of discharge>=1 (1)	Define plan's hospital network if number of discharge>=5 (2)
Panel A: outcome=WTP		
High charge	-0.136	-0.462
	(0.131)	(0.311)
Low charge	0.275**	0.736***
	(0.112)	(0.282)
Superior	0.067	0.317
	(0.196)	(0.411)
Inferior	-0.163	0.361
	(0.157)	(0.355)
Sample size	688	318
Panel B: outcome=probability of inclusion		
WTP	0.029*	-0.001
	(0.015)	(0.016)
Sample size	507	229
Panel C: outcome=probability of exclusion		
WTP	0.005	-0.012
	(0.015)	(0.021)
Sample size	529	259

Notes: Standard errors are reported in parentheses. Each column of each panel is a separate regression. We assume that a hospital is in the network if at least 1 patient is admitted from the particular insurer in column (1); and we assume that a hospital is in the network if at least 5 patients are admitted from the particular insurer in column (5). In panel A, the dependent variable is willingness-to-pay at hospital-plan level. High charge (low charge) is a dummy equal to one if hospital charge is in the top (bottom) quartile in the report card episode. Superior and inferior are defined based on in-hospital mortality ratings. In panel B, the dependent variable is a dummy equal to one if the hospital is newly included into a given network compared to last report card episode. In panel C, the dependent variable is a dummy equal to one if the hospital is dropped in a given network compared to the following report card episode. All specifications include report card episode fixed effects, hospital characteristics, including non-for-profit status, indicator for member of teaching council, doctors per bed, and bed size, and plan fixed effects. Specifications in panel C also include hospital fixed effects. *significant at 10%; ** significant at 5%; ***significant at 1%.

Reference

Aiken, C. E. and S. E. Ozanne, 2013. "Sex Differences in Developmental Programming Models," *Reproduction*, 145 (1), R1–R13.

Aizer, A. and L. Stroud, 2010. "Education, Knowledge and the Evolution of Disparities in Health," *National Bureau of Economic Research*.

Almond, D., 2006. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 US Population," *Journal of Political Economy*, 114 (4), 672–712.

Almond, D., and J. Currie, 2011. "Killing Me Softly: The Fetal Origins Hypothesis," *The Journal of Economic Perspectives*, pp. 153–172.

Alwasel, S. H., A. Harrath, J. S. Aljarallah, Z. Abotalib, C. Osmond, S. Y. Omar, I. Khaled, and D. J. Barker, 2013. "Intergenerational Effects of *In Utero* Exposure to Ramadan in Tunisia," *American Journal of Human Biology*, 25 (3), 341–343.

Anderberg, D., A. Chevalier, and J. Wadsworth, 2011. "Anatomy of a Health Scare: Education, Income and the MMR Controversy in the UK," *Journal of Health Economics*, 30 (3), 515–530.

Barker, D. J., 1990. "The Fetal and Infant Origins of Adult Disease," *British Medical Journal*, 301 (6761), 1111.

Barker, D. J., 1995. "Fetal Origins of Coronary Heart Disease," *British Medical Journal*, 311 (6998), 171–174.

Berenson, R. et al., "The Growing Power of Some Providers to Win Steep Payment Increases from Insurers Suggests Policy Remedies May Be Needed," *Health Affairs*, 31 (5).

Bernard, S. et al., 2001. "Autism: A Novel Form of Mercury Poisoning," *Medical Hypotheses*, 56 (4), 462–471.

Bhalotra, S. and S. Rawlings, 2013. "Gradients of the Intergenerational Transmission of Health in Developing Countries," *Review of Economics and Statistics*, 95 (2), 660–672.

Brown, M. A., G. K. Davis, and L. McHugh, 2001. "The Prevalence and Clinical Significance of Nocturnal Hypertension in Pregnancy," *Journal of Hypertension*, 19 (8), 1437–1444.

Buchbinder, A. et al., 2002. "Adverse Perinatal Outcomes Are Significantly Higher in Severe Gestational Hypertension than in Mild Preeclampsia," *American Journal of Obstetrics and Gynecology*, 186 (1), 66–71.

Clifton, V. L., 2005. "Sexually Dimorphic Effects of Maternal Asthma during Pregnancy on Placental Glucocorticoid Metabolism and Fetal Growth," *Cell and Tissue Research*, 322(1), 63–71.

Collins, J. W., R. J. David, N. G. Prachand, and M. L. Pierce, 2003. "Low Birth Weight across Generations," Maternal and Child Health Journal, 7 (4), 229–237.

Collins, J. W., S.-Y. Wu, and R. J. David, 2002. "Differing Intergenerational Birth Weights among the Descendants of US-Born and Foreign-Born Whites and African Americans in Illinois," *American Journal of Epidemiology*, 155 (3), 210–216.

Commentary, 2008. "Editorial: Rare Conditions," *Akron Beacon Journal*.

Currie, J. and E. Moretti, 2007. "Biology as Destiny? Short-and Long-Run Determinants of Intergenerational Transmission of Birth Weight," *Journal of Labor Economics*, 25 (2).

Darley, J. M. and P. H. Gross, 1983. "A Hypothesis-Confirming Bias in Labeling Effects," *Journal of Personality and Social Psychology*, 44 (1), 20.

DeStefano, F., C. S. Price, and E. S. Weintraub, 2013. "Increasing Exposure to Antibody-Stimulating Proteins and Polysaccharides in Vaccines Is Not Associated with Risk of Autism," *The Journal of Pediatrics*, 163 (2), 561–567.

DeStefano, F. et al., 2004. "Age at First Measles-Mumps-Rubella Vaccination in Children with Autism and School-Matched Control Subjects: A Population-Based Study in Metropolitan Atlanta," *Pediatrics*, 113 (2), 259–266.

Drake, A. J. and B. R. Walker, 2004. "The Intergenerational Effects of Fetal Programming: Non-genomic Mechanisms for the Inheritance of Low Birth Weight and Cardiovascular Risk," *Journal of Endocrinology*, 180 (1), 1–16.

Dranove, D. and A. Sfekas, 2008. "Start Spreading the News: A Structural Estimate of the Effects of New York Hospital Report Cards," *Journal of Health Economics*, 27 (5).

Dranove, D. et al., 2003. "Is More Information Better? The Effects of "Report Cards" on Health Care Providers," *Journal of Political Economy*, 111 (3).

Eriksson, J. G., E. Kajantie, C. Osmond, K. Thornburg, and D. J. Barker, 2010. "Boys Live Dangerously in the Womb," *American Journal of Human Biology*, 22 (3), 330–335.

Esquivel, P. and S. Poindexter, 2014. "Immunity at Risk; More Parents Are Seeking Vaccination Exemptions," *Los Angeles Times*.

Farrington, C., E. Miller, and B. Taylor, 2001. "MMR and Autism: Further Evidence Against a Causal Association," *Vaccine*, 19 (27), 3632–3635.

Foss, K. A. and B. G. Southwell, 2006. "Infant Feeding and the Media: The Relation- ship between Parents' Magazine Content and Breastfeeding, 1972–2000," *International Breastfeeding Journal*, 1 (10), 1–9.

Fraga, M. F. et al., 2005. "Epigenetic Differences Arise during the Lifetime of Monozygotic Twins," *Proceedings of the National Academy of Sciences of the United States of America*, 102 (30), 10604–10609.

Friedrich, J., 1993. "Primary Error Detection and Minimization (PEDMIN) Strategies in Social Cognition: A Reinterpretation of Confirmation Bias Phenomena.," *Psychological Review*, 100 (2), 298.

Genzlinger, N., 2010. "Vaccinations: A Hot Debate Still Burning," *The New York Times*.

Gilbert, P., 1998. "The Evolved Basis and Adaptive Functions of Cognitive Distortions," British *Journal of Medical Psychology*, 71 (4), 447–463.

Gluckman, P. D., M. A. Hanson, and A. S. Beedle, 2007. "Non-genomic Transgenerational Inheritance of Disease Risk," *Bioessays*, 29 (2), 145–154.

Goldenberg, R. L. et al., 2006. "The Alabama Preterm Birth Study: Intrauterine Infection and Placental Histologic Findings in Preterm Births of Males and Females Less than 32 Weeks," *American Journal of Obstetrics and Gynecology*, 195 (6), 1533–1537.

Griffin, D. and A. Tversky, 1992. "The Weighing of Evidence and the Determinants of Confidence," *Cognitive Psychology*, 24 (3), 411–435.

Hales, C. N. and D. J. Barker, 2001. "The Thrifty Phenotype Hypothesis," *British Medical Bulletin*, 60 (1), 5–20.

Hilton, S., M. Petticrew, and K. Hunt, 2007. "Parents' Champions vs. Vested Interests: Who Do Parents Believe about MMR? A Qualitative Study," *BMC Public Health*, 7 (1), 42.

Hilts, P. J., 2000. "House Panel Asks for Study of a Vaccine," *The New York Times*.

Ho, K. and A. Pakes, 2014. "Hospital Choices, Hospital Prices and Financial Incentives to Physicians," *American Economic Review*, 104(12): 3841-84

Ho, K. and R. S. Lee, 2013. "Insurer Competition and Negotiated Hospital Prices," *National Bureau of Economic Research Working Paper*.

Hollier, L. M., D. D. McIntire, and K. J. Leveno, 1999. "Outcome of Twin Pregnancies According to Intrapair Birth Weight Differences," *Obstetrics and Gynecology*, 94 (6), 1006–1010.

Jablonka, E. and M. J. Lamb, 2005. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*, MIT Press.

Katz-field, J., 2011. "Why Immunize?," *Brattleboro Reformer*.

Kayhan, V. O., 2013. "Seeking Health Information on the Web: Positive Hypothesis Testing," *International Journal of Medical Informatics*, 82 (4), 268–275.

Keren, G., 1987. "Facing Uncertainty in the Game of Bridge: A Calibration Study," *Organizational Behavior and Human Decision Processes*, 39 (1), 98–114.

Klebanoff, M. A., N. J. Secher, B. R. Mednick, and C. Schulsinger, 1999. "Maternal Size at Birth and the Development of Hypertension during Pregnancy: A Test of the Barker Hypothesis," *Archives of Internal Medicine*, 159 (14), 1607–1612.

Kleffman, S., 2000. "State's Autism Cases Continue to Increase; Little is Known Why, However One Theory on a Link to Child Vaccinations Ties an International Feud," *Contra Costa Times*.

Kolstad, J., 2013. "Information and Quality When Motivation is Intrinsic: Evidence from Surgeon Report Cards," *American Economic Review*, 103 (7).

Kramer, M. S., 1987. "Determinants of Low Birth Weight: Methodological Assessment and Meta-Analysis.," *Bulletin of the World Health Organization*, 65 (5), 663.

Kunda, Z., 1999. *Social Cognition: Making Sense of People*, MIT press.

Kuzawa, C. W., 2005. "Fetal Origins of Developmental Plasticity: Are Fetal Cues Reliable Predictors of Future Nutritional Environments?," *American Journal of Human Biology*, 17 (1), 5–21.

Kuzawa, C. W. and E. Sweet, 2009. "Epigenetics and the Embodiment of Race: Developmental Origins of US Racial Disparities in Cardiovascular Health," *American Journal of Human Biology*, 21 (1), 2–15.

Lang, U. et al., 2003. "Uterine Blood Flow-A Determinant of Fetal Growth," *European Journal of Obstetrics and Gynecology and Reproductive Biology*, 110, S55–S61.

Levitt, N. S. et al., 2000. "Impaired Glucose Tolerance and Elevated Blood Pressure in Low Birth Weight, Nonobese, Young South African Adults: Early Programming of Cortisol Axis," *The Journal of Clinical Endocrinology and Metabolism*, 85 (12), 4611–4618.

Lindquist, E., 2009. "Family Faces Uncertainty in Dealing with Autism," *The Leader-Telegram (Eau Claire, Wisconsin)*.

Lord, C. G., L. Ross, and M. R. Lepper, 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology*, 37 (11), 2098.

Lucas, A., 1991. "Programming by Early Nutrition in Man," *The Childhood Environment and Adult Disease*, 1991, 38–55.

Lumey, L. H., 1992. "Decreased Birthweights in Infants after Maternal In Utero Exposure to the Dutch Famine of 1944–1945," *Paediatric and Perinatal Epidemiology*, 6 (2), 240–253.

Madsen, K. M. et al., 2002. "A Population-Based Study of Measles, Mumps, and Rubella Vaccination and Autism," *New England Journal of Medicine*, 347 (19), 1477–1482.

Magnus, P., H. K. Gjessing, A. Skrondal, and R. Skjaerven, 2001. "Paternal contribution to birth weight," *Journal of Epidemiology and Community Health*, 55 (12), 873–877.

Mathews, A. and C. MacLeod, 1994. "Cognitive Approaches to Emotion and Emotional Disorders," *Annual Review of Psychology*, 45 (1), 25–50.

Matthews, S. G. and D. I. Phillips, 2010. "Minireview: Transgenerational Inheritance of the Stress Response: A New Frontier in Stress Research," *Endocrinology*, 151 (1), 7–13.

McTernan, C. L. et al., 2001. "Reduced Placental 11β-Hydroxysteroid Dehydrogenase Type 2 mRNA Levels in Human Pregnancies Complicated by Intrauterine Growth Restriction: An Analysis of Possible Mechanisms," *The Journal of Clinical Endocrinology and Metabolism*, 86 (10), 4979–4983.

Meaney, M. J., 2001. "Maternal Care, Gene Expression, and the Transmission of Individual Differences in Stress Reactivity across Generations," *Annual Review of Neuroscience*, 24 (1), 1161–1192.

Mingrone, G. et al., 2008. "Influence of Maternal Obesity on Insulin Sensitivity and Secretion in Offspring," *Diabetes Care*, 31 (9), 1872–1876.

Morgan, H. D., H. G. Sutherland, D. I. Martin, and E. Whitelaw, 1999. "Epigenetic Inheritance at the Agouti Locus in the Mouse," *Nature Genetics*, 23 (3), 314–318.

Morley, R., J. Owens, E. Blair, and T. Dwyer, 2002. "Is Birthweight a Good Marker for Gestational Exposures that Increase the Risk of Adult Disease?," *Paediatric and Perinatal Epidemiology*, 16 (3), 194–199.

Nickerson, R. S., 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," *Review of General Psychology*, 2 (2), 175.

Oswald, M. E. and S. Grosjean, 2004. "4 Confirmation bias," *Cognitive Illusions: A Hand- book on Fallacies and Biases in Thinking, Judgment and Memory*, p. 79.

Park, J. et al., 2013. "Information Valuation and Confirmation Bias in Virtual Communities: Evidence from Stock Message Boards," *Information Systems Research*, 24 (4), 1050–1067.

Peltola, H. et al., 1998. "No Evidence for Measles, Mumps, and Rubella Vaccine-Associated Inflammatory Bowel Disease or Autism in a 14-year Prospective Study," *The Lancet*, 351 (9112), 1327–1328.

Petronis, A., 2010. "Epigenetics as a Unifying Principle in the Aetiology of Complex Traits and Diseases," *Nature*, 465 (7299), 721–727.

Phillips, D. I. et al., 1998. "Elevated Plasma Cortisol Concentrations: A Link between Low Birth Weight and the Insulin Resistance Syndrome?," *The Journal of Clinical Endocrinology and Metabolism*, 83 (3), 757–760.

Price, C. S. et al., 2010. "Prenatal and Infant Exposure to Thimerosal from Vaccines and Immunoglobulins and Risk of Autism," *Pediatrics*, 126 (4), 656–664.

Price, K. C. and C. L. Coe, 2000. "Maternal Constraint on Fetal Growth Patterns in the Rhesus Monkey (Macaca Mulatta): The Intergenerational Link between Mothers and Daughters," *Human Reproduction*, 15 (2), 452–457.

Price, K. C., J. S. Hyde, and C. L. Coe, 1999. "Matrilineal Transmission of Birth Weight in the Rhesus Monkey (Macaca Mulatta) across Several Generations," *Obstetrics and Gynecology*, 94 (1), 128–134.

Rabin, M. and J. L. Schrag, 1999. "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics*, pp. 37–82.

Ramakrishnan, U., R. Martorell, D. G. Schroeder, and R. Flores, 1999. "Role of Intergenerational Effects on Linear Growth," *The Journal of Nutrition*, 129 (2), 544S–549S.

Remmerswaal, D. et al., 2014. "Cognitive Bias in Action: Evidence for a Reciprocal Relation between Confirmation Bias and Fear in Children," *Journal of Behavior Therapy and Experimental Psychiatry*, 45 (1), 26–32.

Reynolds, R. M. et al., 2001. "Altered Control of Cortisol Secretion in Adult Men with Low Birth Weight and Cardiovascular Risk Factors," *The Journal of Clinical Endocrinology and Metabolism*, 86 (1), 245–250.

Richler, J. et al., 2006. "Is There a 'Regressive Phenotype' of Autism Spectrum Disorder Associated with the Measles-Mumps-Rubella Vaccine? A CPEA Study," *Journal of Autism and Developmental Disorders*, 36 (3), 299–316.

Roemer, I., W. Reik, W. Dean, and J. Klose, 1997. "Epigenetic Inheritance in the Mouse," *Current Biology*, 7 (4), 277–280.

Roseboom, T. J. et al., 2001. "Effects of Prenatal Exposure to the Dutch Famine on Adult Disease in Later Life: An Overview," *Molecular and Cellular Endocrinology*, 185 (1), 93–98.

Royer, H., 2009. "Separated at Girth: US Twin Estimates of the Effects of Birth Weight," *American Economic Journal: Applied Economics*, 1 (1), 49–85.

Royer, H. and A. Witman, 2013. "Intergenerational Effects on Health-In Utero and Early Life," *Encyclopedia of Health Economics*.

Simon, D. M. et al., 2006. "Relation of Maternal Low Birth Weight to Infant Growth Retardation and Prematurity," *Maternal and Child Health Journal*, 10 (4), 321–327.

Sinaiko, A. D. and M. B. Rosenthal, 2011. "Increased Price Transparency in Health CareChallenges and Potential Effects," *New England Journal of Medicine*, 364 (10), 891– 894.

Smith, M. J., S. S. Ellenberg, L. M. Bell, and D. M. Rubin, 2008. "Media Coverage of the Measles-Mumps-Rubella Vaccine and Autism Controversy and Its Relationship to MMR Immunization Rates in the United States," *Pediatrics*, 121 (4), e836–e843.

Stark, M. J., I. M. Wright, and V. L. Clifton, 2009. "Sex-Specific Alterations in Placental 11β-Hydroxysteroid Dehydrogenase 2 Activity and Early Postnatal Clinical Course following Antenatal Betamethasone," *American Journal of Physiology-Regulator, yIntegrative and Comparative Physiology*, 297 (2), R510–R514.

Stehr-Green, P. et al., 2003. "Autism and Thimerosal-Containing Vaccines: Lack of Consistent Evidence for an Association," *American Journal of Preventive Medicine*, 25 (2), 101–106.

Stryker, J. E., 2003. "Articles Media and Marijuana: A Longitudinal Analysis of News Media Effects on Adolescents' Marijuana Use and Related Outcomes, 1977-1999*," Journal of Health Communication*, 8 (4), 305–328.

Suen, W., 2004. "The Self-Perpetuation of Biased Beliefs," *The Economic Journal*, 114 (495), 377–396.

Taylor, B. et al., 1999. "Autism and Measles, Mumps, and Rubella Vaccine: No Epidemiological Evidence for A Causal Association," *The Lancet*, 353 (9169), 2026–2029.

Taylor, B. et al., 2002. "Measles, Mumps, and Rubella Vaccination and Bowel Problems or Developmental Regression in Children with Autism: Population Study," *British Medical Journa*l, 324 (7334), 393–396.

Town, R. and G. Vistnes, 2001. "Hospital Competition in HMO Networks," *Journal of Health Economics*, 20 (5).

Traut-Mattausch, E. et al., 2011. "Are There "His" and "Her" Types of Decisions? Exploring Gender Differences in the Confirmation Bias," *Sex Roles*, 65 (3-4), 223–233.

Trivers, R. L. and D. E. Willard, 1973. "Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring," *Science*, 179 (4068), 90–92.

Verstraeten, T. et al., 2003. "Safety of Thimerosal-Containing Vaccines: A Two-Phased Study of Computerized Health Maintenance Organization Databases," *Pediatrics*, 112 (5), 1039–1048.

Wakefield, A. J. et al., 1998. "RETRACTED: Ileal-lymphoid-nodular Hyperplasia, Non-Specific colitis, and Pervasive Developmental Disorder in Children," *The Lancet*, 351 (9103), 637–641.

Wang, J., J. Hockenberry, S.-Y. Chou, and M. Yang, 2011. "Do Bad Report Cards Have Consequences? Impacts of Publicly Reported Provider Quality Information on the CABG Market in Pennsylvania," *Journal of Health Economics*, 30 (2), 392–407.

Weaver, I. C. et al., 2004. "Epigenetic Programming by Maternal Behavior," *Nature Neuroscience*, 7 (8), 847–854.

Worthman, C. M. and J. Kuzara, 2005. "Life History and the Early Origins of Health Differentials," *American Journal of Human Biology*, 17 (1), 95–112.

Wright, J. A. and C. Polack, 2006. "Understanding Variation in Measles-Mumps-Rubella Immunization CoverageA Population-Based Study," *The European Journal of Public Health*, 16 (2), 137–142.

Yanovitzky, I. and J. Stryker, 2001. "Mass Media, Social Norms, and Health Promotion Efforts a Longitudinal Study of Media Effects on Youth Binge Drinking," *Communication Research*, 28 (2), 208–239.

Zaren, B., G. Lindmark, and L. Bakketeig, 2000. "Maternal Smoking Affects Fetal Growth More in the Male Fetus," *Paediatric and Perinatal Epidemiology*, 14 (2), 118–126.

Appendix A

Table A1 Estimated IUGR Thresholds for G2 and G3 Singleton

	The third generation: 1999-2006			The second generation: 1979-1985		
Gestation (weeks)	Observation	SGA (5^{th} pctl.)	2SD < mean	Observation	SGA (5^{th} pctl.)	2SD < mean
31	2,449	1,100	875	1,550	1,250	646
32	3,781	1,280	1,127	5,669	1,400	1,086
33	5,577	1,450	1,285	3,047	1,500	1,216
34	9,899	1,665	1,501	6,206	1,600	1,319
35	20,625	1,920	1,718	8,046	1,775	1,497
36	59,198	2,160	1,956	31,227	2,000	1,669
37	185,270	2,400	2,206	33,865	2,200	1,989
38	428,295	2,550	2,371	156,372	2,400	2,236
39	462,182	2,650	2,472	209,697	2,600	2,407
40	339,773	2,700	2,550	1,966,796	2,700	2,469
41	74,393	2,780	2,605	98,115	2,750	2,555
42	6,793	2,750	2,575	55,538	2,750	2,514
43	579	2,680	2,436	11,762	2,700	2,505
44	146	2,600	2,394	3,942	2,700	2,462
45	75	2,290	2,416	1,056	2,700	2,494

Notes: The estimation sample for the third generation includes all singleton birth in Taiwan between 1999 and 2006 with gestation between 31 and 45 weeks and birth weight between 400 and 6500 grams. The estimation sample for the second generation includes all singleton births in Taiwan between 1978 and 1985 with gestation between 31 and 45 weeks and birth weight between 400 and 6,500 grams.

Table A2 Effect of Mother's LBW on Child's IUGR[1]

	(1)	(2)	(3)	(4)	(5)	(6)
Sample: gestation ≥ 31 weeks (N=280,030)						
Impact of G2 LBW on G3 SGA (5th pctl.)	0.0519***	0.0523***	0.0523***	0.0518***	0.0516***	0.0240***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.009)
Impact of G2 LBW on G3 2SD < mean	0.0230***	0.0231***	0.0231***	0.0228***	0.0227***	0.0147***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.006)
Sample: gestation between 37 and 42 weeks (N=261,478)						
Impact of G2 LBW on G3 FT LBW	0.0370***	0.0370***	0.0370***	0.0365***	0.0364***	0.0195***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)
G3 gestational age		Y	Y	Y	Y	Y
Pre-determined variables before G3 birth[2]			Y	Y	Y	Y
Jointly-determined variables with pregnancy of G3[3]				Y	Y	Y
G2 spousal characteristics[4]					Y	Y
G1 (grandmother) fixed effects						Y

[1] Standard errors are clustered at child's hospital and year level in parentheses. Each column of each row is a separate regression. All dependent variables are dummies for G3's IUGR. Coefficients are reported for indicators of G2's LBW. *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

[2] Pre-determined variables before G3 birth include G3's gender, birth year dummies, G2's birth order dummies (first, second, and third), birth place dummies (hospital and clinics or maternity homes), G1's marital status at G2's birth, and interactions between dummies for G1's county of residence at the time of G2's birth and dummies for G2's birth year.

[3] Jointly-determined variables with pregnancy of G3 include G3's birth order dummies (first, second, and third), G2's five education dummies, two age dummies and marital status at child's birth, town-level average income at child's birth year, and interactions between dummies for G2's county and dummies for G3's birth year.

[4] G2 spousal characteristics include five dummies for spousal education and six age dummies.

Table A3 Effect of Father's LBW on Child's IUGR[1]

	(1)	(2)	(3)	(4)	(5)	(6)
Sample: gestation ≥ 31 weeks (N=125,078)						
Impact of G2 LBW on G3 SGA (5th pctl.)	0.0275***	0.0276***	0.0271***	0.0266***	0.0263***	0.0167
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.024)
Impact of G2 LBW on G3 2SD < mean	0.0134***	0.0134***	0.0131***	0.0129***	0.0128***	0.0248*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.015)
Sample: gestation between 37 and 42 weeks (N=116,509)						
Impact of G2 LBW on G3 FT LBW	0.0206***	0.0206***	0.0201***	0.0196***	0.0194***	0.0110
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.019)
G3 gestational age		Y	Y	Y	Y	Y
Pre-determined variables before G3 birth[2]			Y	Y	Y	Y
Jointly-determined variables with pregnancy of G3[3]				Y	Y	Y
G2 spousal characteristics[4]					Y	Y
G1 (grandmother) fixed effects						Y

[1] Standard errors are clustered at child's hospital and year level in parentheses. Each column of each row is a separate regression. All dependent variables are dummies for G3's IUGR. Coefficients are reported for indicators of G2's LBW. *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

[2] Pre-determined variables before G3 birth include G3's gender, birth year dummies, G2's birth order dummies (first, second, and third), birth place dummies (hospital and clinics or maternity homes), G1's marital status at G2's birth, and interactions between dummies for G1's county of residence at the time of G2's birth and dummies for G2's birth year.

[3] Jointly-determined variables with pregnancy of G3 include G3's birth order dummies (first, second, and third), G2's five education dummies, two age dummies and marital status at child's birth, town-level average income at child's birth year, and interactions between dummies for G2's county and dummies for G3's birth year.

[4] G2 spousal characteristics include five dummies for spousal education and five age dummies.

Table A4 Variations in Maternal and Paternal Impacts by Family Size

	Dependent variables			
	LBW	SGA (5^{th} pctl.)	2SD < mean	FT LBW
	(1)	(2)	(3)	(4)
Panel A: the maternal sample				
G2 LBW (or IUGR)	0.0177*	0.0231***	0.0204**	0.0228**
	(0.010)	(0.008)	(0.008)	(0.010)
G2 LBW (or IUGR) x large family	0.0151	-0.0020	-0.0039	-0.0026
	(0.016)	(0.014)	(0.014)	(0.017)
Sample size	280,030	280,030	280,030	255,100
Panel B: the paternal sample				
G2 LBW (or IUGR)	-0.0206	0.0238	0.0061	0.0228
	(0.025)	(0.024)	(0.022)	(0.028)
G2 LBW (or IUGR) x large family	0.0309	-0.0455	0.0084	0.0229
	(0.046)	(0.041)	(0.039)	(0.050)
Sample size	125,078	125,078	125,078	113,639

Notes: Standard errors clustered at child's hospital and year level are reported in parentheses. Each column of each panel is a separate regression. Results in Panel A are estimated using the maternal sample; and results in Panel B are estimated using the paternal sample. The dependent variables are dummies for child's LBW (or IUGR). Coefficients are reported for G2's LBW (or corresponding IUGR indicator) and its interaction with an indicator for large family size. Large family is a dummy equal to one if the child has at least two older siblings. All regressions include the full set of control variables and grandmother fixed effects. *** Significant at the 1 percent level; ** Significant at the 5 percent level; * Significant at the 10 percent level.

Table A5 Test on Identical Effect of G2 LBW on the Probability of Being Observed in the Samples by Group

	p-values of F-test	
	G2 females	G2 males
	(1)	(2)
G2 birth cohort	0.708	0.020
G1 marital status at G2 birth	0.662	0.752
G2 birth order	0.123	0.177
G1 age at G2 birth	0.142	0.457
G2 birth place	0.476	0.040
G1 years of schooling at G2 birth	0.888	0.748
G1 county of residence at G2 birth	0.837	0.858
G1 spousal age at G2 birth	0.001	0.049

Notes: P-values of F-test on null hypothesis that the effect of LBW on the probability of being observed in the maternal (paternal) sample is the same across each of the above characteristics are reported for G2 females (males). The estimation sample includes all female (in column (1)) or male (in column (2)) singleton births in Taiwan between 1978 and 1985 with gestation between 31 and 45 weeks and birth weight between 400 and 6,500 grams. The probability of being observed is the probability that the singleton mother or father is observed giving birth to a singleton in Taiwan between 1999 and 2006.

Table A6 Test on Identical Effect of G2 LBW (or IUGR) on Child's LBW (or IUGR) by Group

	Dependent variables			
	LBW	SGA (5^{th} pctl.)	2SD < mean	FT LBW
	(1)	(2)	(3)	(4)
Panel A: the maternal sample				
G1 spousal age at G2 birth	0.357	0.418	0.754	0.174
Panel B: the paternal sample				
G2 birth cohort	0.128	0.239	0.357	0.541
G2 birth place	0.955	0.022	0.479	0.703
G1 spousal age at G2 birth	0.716	0.046	0.903	0.243

Notes: P-values of F-test on null hypothesis that the effect of G2's LBW (or IUGR) on child's LBW (or IUGR) is the same across variables that fail the test presented in Table A5. Results in panel A are estimated using the maternal sample; and results in panel B are estimated using the paternal sample. Each column of each row is a separate regression. The dependent variables are dummies for child's LBW (or IUGR). Interactions between corresponding dummies for mother's (or father's) LBW (or IUGR) and each categories of the presented variables are included in the regression. All regressions also include the full set of control variables and grandmother fixed effects.

Table A7 Gender Differences in Maternal Impacts of LBW (or IUGR) by Family Size

	Dependent variables			
	LBW	SGA (5^{th} pctl.)	2SD < mean	FT LBW
	(1)	(2)	(3)	(4)
G2 LBW (or IUGR)	0.0261***	0.0384***	0.0255***	0.0305***
	(0.009)	(0.008)	(0.007)	(0.009)
G2 LBW (or IUGR) x male	-0.0056	-0.0305***	-0.0145**	-0.0177**
	(0.007)	(0.006)	(0.006)	(0.007)
G2 LBW (or IUGR) x male x large family	-0.0104	0.0044	0.0346**	0.0125
	(0.016)	(0.014)	(0.014)	(0.018)
Sample size	280,030	280,030	280,030	255,100

Notes: Standard errors clustered at child's hospital and year level are reported in parentheses. Each column is a separate regression. The dependent variables are dummies for child's LBW (or IUGR). Coefficients are reported for corresponding dummy for mother's LBW (or IUGR), its interaction with indicator for male birth, and a triple interaction term among mother's LBW (or corresponding IUGR indicator), an indicator for male birth, and an indicator for large family size. Large family is a dummy equal to one if the child has at least two older siblings. All regressions include the full set of control variables and grandmother fixed effects.

Table A8 Probability of Observing a Male Birth as a Function of Mother's LBW (or IUGR)

	Independent variables			
Dependent variable	G2 LBW	G2 SGA (5^{th} pctl.)	G2 2SD < mean	G2 FT LBW
	(1)	(2)	(3)	(4)
Male	0.0077	0.0144	0.0013	0.0035
	(0.020)	(0.016)	(0.024)	(0.025)
Sample size	280,030	280,030	280,030	255,100

Notes: Standard errors clustered at child's hospital and year level are reported in parentheses. Each column of each panel is a separate regression. The dependent variable is an indicator equal to one if the child is a male. Coefficients are reported for dummies for mother's (or father's) LBW (or IUGR). Results are estimated using the maternal sample. All regressions include the full set of control variables and grandmother fixed effects.

Table A9 Variations in the Impact of Mother's LBW (or IUGR) on Male Birth by SES

	High SES group			
	County-level unemployment rate	Town-level average income	County-level parental education	County-level educational improvement
	(1)	(2)	(3)	(4)
Panel A: impact of G2 LBW on G3 male birth				
G2 x LBW	0.0062	0.0071	-0.0026	0.0022
	(0.022)	(0.022)	(0.022)	(0.022)
G2 LBW x high SES	0.0027	0.0016	0.0274	0.0121
	(0.016)	(0.028)	(0.022)	(0.022)
Panel B: impact of G2 SGA (5th pctl.) on G3 male birth				
G2 SGA (5th pctl.)	0.0119	0.0119	0.0117	0.0033
	(0.017)	(0.018)	(0.017)	(0.018)
G2 SGA (5th pctl.)x high SES	0.0047	0.0067	0.0072	0.0237
	(0.013)	(0.022)	(0.018)	(0.017)
Panel C: impact of G2 2SD < mean on G3 male birth				
G2 2SD < mean	-0.0080	-0.0038	-0.0045	-0.0071
	(0.027)	(0.027)	(0.026)	(0.027)
G2 2SD < mean x high SES	0.0168	0.0145	0.0161	0.0180
	(0.020)	(0.035)	(0.028)	(0.027)
Sample size	280,030	280,030	280,030	280,030
Panel D: impact of G2 FT LBW on G3 male birth				
G2 FT LBW	-0.0003	0.0067	0.0013	0.0004
	(0.027)	(0.028)	(0.027)	(0.028)
G2 FT LBW x high SES	0.0068	-0.0088	0.0060	0.0068
	(0.021)	(0.036)	(0.028)	(0.028)
Sample size	255,100	255,100	255,100	255,100

Notes: Standard errors clustered at child's hospital and year level are reported in parentheses. Each column of each panel is a separate regression. The dependent variable is a dummy equal to one if the child is a male. Coefficients are reported for mother's LBW (or IUGR) and the interaction between mother's LBW (or IUGR) and indicators for high SES group. In column (1), high SES is defined as county-level unemployment rate at child's birth year below the mean of the sample; in column (2), high SES is defined as town-level average income at child's birth year above the mean of the sample; in column (3), high SES is defined as county-level percentage of at least one parent with years of schooling greater than 12 at child's birth above the mean of the sample; and in column (4), high SES is defined as change in county-level average parental education when giving birth from G1 to G2 is above the mean of the sample.

Appendix B

Following the conceptual framework of "fetal origins" put forward by Douglas (2006), we demonstrate that the greater adaptability to intrauterine environment of females indicated by the literature (Aiken and Ozanne 2013) leads to a higher observed correlation with maternal LBW (or IUGR) because of lower mortality rates. We assume that there are two types of fetuses in the population of both genders: 1) those with non-LBW (or non-IUGR) mothers, who are not likely to be affected by the intergenerational fetal programming effect; and 2) those with LBW (or IUGR) mothers, who are likely to inherit the fetal programming effect. For simplicity, we assume further that the underlying health distributions for male and female fetuses are identical for both types. Let h be the unobserved health of the individual. For those potentially affected by fetal programming transmission, the underlying health of both genders will deteriorate, causing a left shift in both distributions by the same amount, c.

For a normal fetus, if h falls below a survival threshold, d, then fetal mortality occurs, so we are unable to observe the individual. The fetus will be identified as LBW (or IUGR) later, at birth, if $d < h \leq a$, where a is a fixed threshold determined by the health distribution of the entire population (including both types). Given these thresholds, the fetal mortality rate, $F(d)$, and LBW (or IUGR) incidence, $\frac{F(a)-F(d)}{1-F(d)}$, are the same across gender, where $F(.)$ is the cumulative distribution function.

For a fetus more likely to be affected by intergenerational fetal programming, we consider an extreme case where females are completely adaptable while males are completely inadaptable to the potential effect. Then, the survival threshold will also decrease to $d - c$ for a female fetus of this type, resulting in an unchanged fetal mortality rate. However, the survival threshold will remain constant for a male fetus of this type, leading to a higher fetal mortality rate. Thus, the

LBW (or IUGR) incidence, given that the fetus has a LBW (or IUGR) mother, will be $\frac{F(a+c)-F(d)}{1-F(d)}$ for females and $\frac{F(a+c)-F(d)}{1-F(d+c)}$ for males.

The change in LBW (or IUGR) incidence between the two types by each gender corresponds to the observed intergenerational correlation in LBW (or IUGR). Because the baseline incidence is the same for males and females, we need only compare the incidence for a fetus with a LBW (or IUGR) mother for males and females. We can show that $\frac{F(a+c)-F(d)}{1-F(d)} > \frac{F(a+c)-F(d)}{1-F(d+c)}$, which means intergenerational correlation among females is stronger than among males.

Although a different sensitivity to the intergenerational fetal programming effect by gender can affect this gender difference, the proof clearly shows that there is a possibility that stronger intergenerational correlation in LBW (or IUGR) in females can be observed even when the underlying health distributions of both genders are affected to the same extent. Therefore, a robustness check for the effect of mother's LBW (or IUGR) on sex ratio is necessary.

Appendix C

Table C1 Mechanisms for Biased Beliefs: Single Information Exposures—Raw Values

	(1)	(2)	(3)	(4)
College x ln(year)	0.0121*	0.0146**	0.0155**	0.0154**
	(0.006)	(0.006)	(0.006)	(0.006)
College x post	0.0014	0.0008	0.0006	0.0011
	(0.005)	(0.005)	(0.005)	(0.005)
College x ln(year) x autism	0.0008	0.0023	0.0035	0.0061
	(0.006)	(0.006)	(0.006)	(0.006)
College x ln(year) x reported cases	-0.0017***	-0.0014**	-0.0014**	-0.0016**
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x news counts	0.0001	0.0001	0.0001	0.0001
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x news for vaccine	0.0032	0.0031	0.0027	0.0032
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ science	-0.0025	-0.0019	-0.0017	-0.0024
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ authority	-0.0034**	-0.0032**	-0.0030**	-0.0031**
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x news w/ parents	0.0023	0.0021	0.0020	0.0022
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x autism search	0.0008	0.0001	-0.0001	-0.0003
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x outbreaks search	-0.0007***	-0.0007***	-0.0007***	-0.0008***
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x "vaccine autism" search	0.0053***	0.0045***	0.0045***	0.0044***
	(0.001)	(0.001)	(0.001)	(0.001)
Sample size	271,478	271,478	271,478	271,478
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Coefficients are reported for triple interaction terms using all the information exposures presented in Table 3. Raw values are used directly for each type of information exposure. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table C2 Mechanisms for Biased Beliefs: Information Sources--News Count as a Safe Signal

	(1)	(2)	(3)	(4)
College x ln(year)	0.0117***	0.0118***	0.0121***	0.0119***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0014	0.0008	0.0007	0.0011
	(0.005)	(0.005)	(0.005)	(0.005)
Disease Prevalence rate	0.0001	0.0002	0.0003	0.0008
College x ln(year) x harmful signal	(0.001)	(0.001)	(0.001)	(0.001)
	-0.0022***	-0.0019**	-0.0020**	-0.0021**
College x ln(year) x safe signal	(0.001)	(0.001)	(0.001)	(0.001)
News coverage	0.0003	0.0001	0.0001	0.0003
College x ln(year) x harmful signal	(0.001)	(0.001)	(0.001)	(0.001)
	0.0004	0.0008	0.0006	0.0011
College x ln(year) x safe signal	(0.002)	(0.002)	(0.002)	(0.002)
Online searches	0.0053***	0.0043***	0.0042***	0.0037**
College x ln(year) x harmful signal	(0.001)	(0.001)	(0.001)	(0.001)
	-0.0015	-0.0013	-0.0014	-0.0017
College x ln(year) x safe signal	(0.001)	(0.001)	(0.001)	(0.001)
Sample size	271,478	271,478	271,478	271,478
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Coefficient of the triple interaction term using harmful signal of disease prevalence rate is multiplied by 100 for reporting purpose. For disease prevalence rate, the harmful signal is the z-score of autism prevalence rate; and the safe signal is the z-score of reported total cases of measles, mumps, and rubella. For news coverage, the harmful signal is the z-scores of percentage of news with parents' opinions; and the safe signal is the average z-scores of news count, percentage of news encouraging immunization, with words from authorities, and scientific proofs. For web searches, the harmful signal is the z-score of search index for measles, mumps, and rubella; and the safe signal is the average of z-scores of search index for autism and "vaccine and autism" topics. All controls in Table 4 are included. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table C3 Mechanisms for Biased Beliefs: Information Attitudes--News Count as a Safe Signal

	(1)	(2)	(3)	(4)
College x ln(year)	0.0111***	0.0112***	0.0115***	0.0112***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0014	0.0009	0.0007	0.0011
	(0.005)	(0.005)	(0.005)	(0.005)
College x ln(year) x harmful signal	0.0061***	0.0048***	0.0049***	0.0052***
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x safe signal	-0.0006	-0.0002	-0.0005	-0.0004
	(0.001)	(0.001)	(0.002)	(0.002)
Sample size	271,478	271,478	271,478	271,478
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Harmful composite is the average z-scores of autism prevalence rate, percentage of news with parents' opinions, and search index for autism, and "autism and vaccine" topic. Safe composite is the average z-scores of news count, reported total cases for measles, mumps, and rubella, percentage of news encouraging immunization, with words from authority, and scientific proofs, and search index for measles, mumps, and rubella. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table C4 Impact of Biased Beliefs—Non-Hispanic Black and White

	(1)	(2)	(3)	(4)
College x ln(year)	0.0125***	0.0119***	0.0121***	0.0121***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0010	0.0007	0.0007	0.0008
	(0.006)	(0.006)	(0.006)	(0.006)
Sample size	196,428	196,428	196,428	196,428
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The sample is restricted to non-Hispanic Black and White children. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. All controls in Table 4 are included. No area specific time trend is controlled in columns (1) and (2). We include region specific year effects in columns (3) and (4), division specific year effects in columns (5) and (6), and state specific log linear time trend in columns (7) and (8). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table C5 Mechanisms for Biased Beliefs: Single Information Exposures--Non-Hispanic Black and White

	(1)	(2)	(3)	(4)
College x ln(year)	0.0120***	0.0118***	0.0120***	0.0119***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	-0.0011	-0.0009	-0.0007	-0.0003
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x autism	-0.0029***	-0.0025***	-0.0025***	-0.0029***
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x reported cases	0.0013	0.0007	0.0007	0.0014
	(0.002)	(0.002)	(0.002)	(0.003)
College x ln(year) x news counts	0.0029	0.0029	0.0021	0.0031
	(0.003)	(0.003)	(0.003)	(0.004)
College x ln(year) x news for vaccine	-0.0018	-0.0011	-0.0006	-0.0014
	(0.002)	(0.002)	(0.002)	(0.003)
College x ln(year) x news w/ science	-0.0040*	-0.0038*	-0.0036*	-0.0038*
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ authority	0.0016	0.0015	0.0013	0.0017
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x news w/ parents	0.0027**	0.0020	0.0020	0.0018
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x autism search	-0.0037***	-0.0035***	-0.0035***	-0.0042***
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x outbreaks search	0.0039***	0.0033***	0.0033***	0.0031***
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x “vaccine autism” search	0.0010	0.0007	0.0008	0.0008
	(0.006)	(0.006)	(0.006)	(0.006)
Sample size	196,428	196,428	196,428	196,428
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The sample is restricted to non-Hispanic Black and White children. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Coefficients are reported for triple interaction terms using all the information exposures presented in Table 3. We use z-scores for each type of information exposure. All controls in Table 4 are included. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table C6 Mechanisms for Biased Beliefs: Information Sources--Non-Hispanic Black and White

	(1)	(2)	(3)	(4)
College x ln(year)	0.0112***	0.0110***	0.0112***	0.0110***
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0010	0.0008	0.0008	0.0008
	(0.006)	(0.006)	(0.006)	(0.006)
Disease Prevalence rate				
College x ln(year) x harmful signal	-0.0013	-0.0012	-0.0010	-0.0005
	(0.001)	(0.001)	(0.001)	(0.001)
College x ln(year) x safe signal	-0.0030***	-0.0026***	-0.0027***	-0.0031***
	(0.001)	(0.001)	(0.001)	(0.001)
News coverage				
College x ln(year) x harmful signal	0.0021	0.0016	0.0016	0.0024
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x safe signal	-0.0021	-0.0015	-0.0017	-0.0013
	(0.002)	(0.002)	(0.002)	(0.002)
Web searches				
College x ln(year) x harmful signal	0.0073***	0.0063***	0.0063***	0.0058***
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x safe signal	-0.0024*	-0.0022*	-0.0022*	-0.0028**
	(0.001)	(0.001)	(0.001)	(0.001)
Sample size	196,428	196,428	196,428	196,428
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The sample is restricted to non-Hispanic Black and White children. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. For disease prevalence rate, the harmful signal is the z-score of autism prevalence rate; and the safe signal is the z-score of reported total cases of measles, mumps, and rubella. For news coverage, the harmful signal is the average z-scores of news count and percentage of news with parents' opinions; and the safe signal is the average z-scores of percentage of news encouraging immunization, with words from authorities, and scientific proofs. For web searches, the harmful signal is the z-score of search index for measles, mumps, and rubella; and the safe signal is the average of z-scores of search index for autism and "vaccine and autism" topics. All controls in Table 4 are included. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table C7 Mechanisms for Biased Beliefs: Information Attitudes—Non-Hispanic Black and White

	(1)	(2)	(3)	(4)
College x ln(year)	0.0105***	0.0103**	0.0105***	0.0103**
	(0.004)	(0.004)	(0.004)	(0.004)
College x post	0.0011	0.0008	0.0008	0.0008
	(0.006)	(0.006)	(0.006)	(0.006)
College x ln(year) x composite harmful signal	0.0071***	0.0057***	0.0060***	0.0070***
	(0.002)	(0.002)	(0.002)	(0.002)
College x ln(year) x composite safe signal	-0.0028	-0.0020	-0.0025	-0.0030
	(0.002)	(0.002)	(0.002)	(0.002)
Sample size	196,428	196,428	196,428	196,428
Area specific time trend	No	Region x year FE	Division x year FE	State x ln(year)

Notes: The Sample is restricted to non-Hispanic Black and White children. The outcome variable is a dummy equal to one if MMR shot is not up-to-date for the child. Harmful composite is the average z-scores of autism prevalence rate, news count, percentage of news with parents' opinions, and search index for autism, and "autism and vaccine" topic. Safe composite is the average z-scores of reported total cases for measles, mumps, and rubella, percentage of news encouraging immunization, with words from authority, and scientific proofs, and search index for measles, mumps, and rubella. No area specific time trend is controlled in columns (1). We include region specific year effects in columns (2), division specific year effects in columns (3), and state specific log linear time trend in columns (4). Standard errors are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Biography

Mengcen Qian was born in Shanghai, China. She attended the School of Management at Fudan University where she earned her bachelor degree in Financial Management in 2010. During her undergraduate studies, she minored in Biotechnology at East China University of Science and Technology.

She continued her graduate studies in the doctoral program in Economics at Lehigh University in the Fall of 2010. She was supported by teaching and research assistantships, and by Warren-York dissertation fellowship awarded by the Department of Economics. Her fields of specialization include health economics, applied econometrics, and applied microeconomics.