TOWARD A BROADER EXPLANATION OF COMPLIANCE AND NONCOMPLIANCE IN CHILDHOOD IMMUNIZATION

by

YONGCHONG MAO

(Under the Direction of Tim Park and Michael Wetzstein)

ABSTRACT

An economic analysis is presented on the factors that affect household's demand for childhood immunization. The analysis is mainly based on the National Immunization Survey (NIS) 2003 Public-Use Data File and supplemented with state Medicaid variables and influenza level variables. A logit regression is estimated for the up-to-date immunization status of children. The regression results imply that factors such as age of children, income, mother's age and education, and Medicaid program have a positive influence on household's demand for children's vaccine. In contrast, the number of children and comprehensive care providers in the household are negatively correlated with up-to-date vaccine status.

Oaxaca-Blinder decomposition technique is used in the research to indicate: How much of the gap between vaccination status of two groups of different mother's education is due to difference in characteristics and difference in the estimated coefficients. The decomposition results imply that approximately 35% of the education effect on up-to-date vaccine status can be explained by a household's immunization preference.

INDEX WORDS: Childhood Immunization, Vaccination, CDC, NIS, Oaxaca-Blinder Decomposition, Influenza.

TOWARD A BROADER EXPLANATION OF COMPLIANCE AND NONCOMPLIANCE IN CHILDHOOD IMMUNIZATION

by

YONGCHONG MAO

Bachelor of Economics, East China University of Science and Technology, China, 2003

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2005

TOWARD A BROADER EXPLANATION OF COMPLIANCE AND NONCOMPLIANCE IN CHILDHOOD IMMUNIZATION

by

YONGCHONG MAO

Major Professors: Tim Park
Michael Wetzstein

Committee: Mark Messonnier
Charlie Huang

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2005

Dedicated To Those Who Love Me and Care for Me

ACKNOWLEDGEMENTS

The completion of this thesis is the direct outgrowth of my major professors' guidance and my effort. First of all, I would like to express my appreciation to my major professors Dr. Tim Park and Dr. Michael Wetzstein, whose dedication in my research and education I appreciate the most. Not only do they give me advice, guidance and encouragement during my research study and thesis writing but also they boost my interest in health economics. As a student from China, I am greatly affected by Confucius's principles which emphasis that one should always love and respect his teachers particularly those who gave him a lot of help. Thus, all the words in the world would not be enough to express my gratitude to them.

I want to express special thanks to other members of my graduate committee: Dr. Mark Messonnier and Dr. Charlie Huang. Their friendly attitude, encouragement and advice were very crucial to me. Likewise, I would like to thank other members in our department. In particular, our department head, Dr. Fred White provided me with the valuable assistantship. I cannot help not to mention how thankful I am to Chris Peters, who always helps me with computer problems, Jo Anne Norris and Laura Alfonso, who gave me a lot of valuable information. In like manner, I want to thank all of my officemates and my graduate student fellows in our department for their friendship, encouragement and strong support.

Finally, the deepest gratitude and love goes to my family: my parents, who brought me up and offer constant financial and spiritual support during my stay in Athens, Georgia. My dear wife, Xiu Yu, who always makes me company and supports me and loves me and cares me so much. UGA in my mind…

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Problem Statement

The market for vaccines is characterized as an externality. One individual receiving a vaccination is prevented from being a vehicle for possibly infecting others. Numerous other individuals' welfares are enhanced by this individual's choice of being vaccinated or not. Efforts directed toward accounting for this externality require some type of governmental mechanism design for intervention into the vaccine market. By requiring children to be vaccinated, quantity demanded is set by various government agencies. Over half of all childhood vaccines are dispensed through public programs and generally purchased on terms set by the Centers for Disease Control (McGuire).

Vaccines as biologic agents have the properties of relatively high storage costs and are perishable. Production is characterized by economies to scale and augmenting production capacity is costly. Determining the efficient scale of production requires knowledge on the vaccine's demand. With this knowledge an effective pricing policy may then be developed. Unfortunately, the externalities associated with vaccines do not provide a market which reveals individuals' demand for a vaccine.

The emphasis of previous work designed to improve childhood immunization rates has focused on economic issues in the supply of vaccines. However, a recent report indicates that demand should also be highlighted as the key component influencing low immunization rates for preschool children rather than supply constraints or shortages. Poor sales have prompted Glaxo Smith Kline (GSK) to withdraw its Lyme disease vaccine, LYMErix, from the United States market. Lawsuits relating to adverse effects of the vaccine have been filed since its introduction, but the US Food and Drug Administration (FDA) found no proof that LYMErix is dangerous. It appears that lack of demand, not safety concerns, is the reason for the withdrawal.

Despite their enormous impact on global health, preventative vaccination in the major diseases is either non-existent or outdated. The increased incidence and geographical spread of diseases coupled with elevated international travel, urbanization, and the threat of bio-terrorism demands for increased levels of preventative action.

Stephen Cochi, acting director of the Centers for Disease Control and Prevention's National Immunization Program states the major concerns associated with childhood vaccinations are eliminating vaccine shortages, reducing racial and ethnic disparities, and addressing the generally unfounded fears about vaccine safety (Manning). Parental fear of childhood vaccinations is often cited as a reason for childhood immunization noncompliance. Even though such fears are generally unfounded, their persistence among parents may significantly affect childhood immunization compliance and the impact of these fears may differ with respect to key sociodemographic factors. Households of unvaccinated children have preferences toward accepting links to autism, bowel disorders, and other possible negative effects of vaccines that may deter them from getting immunizations for their children. Underlying this fear is parental preferences for being in compliance versus noncompliance. A relatively weak preference for being in compliance may account for low compliance rates. These weak preferences may be the result of unfounded fears coupled with a lack of understanding in the benefits of compliance. Any decline in immunization compliance could result in children vulnerable to infection.

Pauly and Cleff investigate the economics of vaccine policy. Mother's age, education, race, income and child's age impact the probability a child is in compliance. Such factors can be readily exacted from surveys. However, household's preferences for compliance are not observable and difficult to elicit from such surveys. Given this difficulty, household's preferences have not been adequately addressed in current research. However, they are potentially a major determinant in compliance. The importance of increased education addressing the benefits and cost of childhood vaccinations is recognized in a number of studies. Unfortunately, estimates do not currently exist on the degree parental preferences influence their compliance. Without adequate measures of these preferences policies and programs directed toward immunization compliance may not be targeting the right incentives for parental compliance.

Objectives

The overall objective of this research project is to develop a methodology for measuring the magnitude of households' underlying preferences toward immunization compliance. The goal is to determine not only if such preferences influence compliance, but if these preferences exist to measure the degree of influence. Specific objectives are:

1. Based a literature review and a theoretical model for immunization compliance, determine the major factors that may affect households' compliance.
2. Given the National Immunization Survey 2003 as a data base, estimate the probability a household will be in compliance.

3. From this estimation, determine the percentage of the difference in immunization compliance caused by households' preferences toward compliance.

Procedures

A theoretical model for immunization compliance, based on previous research, will provide a foundation for empirically estimating the relations of parents' probability of being in compliance or noncompliance. A logit regression will then be employed using the National Immunization Survey to predict the probability of compliance. Marginal effects of the explanatory variable will be calculated to determine the influence each variable has on this probability. The results from this logit model will then be employed in an Oaxaca-Blinder decomposition of compliance. The Oaxaca-Blinder approach will determine the percentage of the difference in immunization compliance caused by parents' preferences toward compliance. This analysis will determine the magnitude of parents' biases toward immunization on their probability of being in compliance.

CHAPTER 2

REVIEW OF RELATED LITERATURE

Kenkel (1990) made an empirical investigation of consumer health information. Using a new direct measure of information, he treats both information and physician visits as endogenous variables when estimating the demand for medical care. The results show that information increases the probability that a consumer uses medical care, but that conditional on use the quantity of care consumed is not related to information. The results contradict specific implications of models where physicians can create or induce demand for their own services. Probit results suggest that poorly informed consumers tend to underestimate the productivity of medical care in treating illness, which is consistent to our common sense that more informed consumers are significantly more likely to visit a physician. Since Kenkel was the first to explore the determinants of health information, the empirical results are especially attractive and interesting. As expected, education and occupation in a health field are associated with more information. The estimated effect of income implies that heath information is a normal good.

Kenkel’s medical care demand model and his making health information as an endogenous variable play a significant role in health economics. His results for the other explanatory variables in the medical care demand model generally conform to prior expectations. Insurance coverage for physician visits increase demand. Hospital coverage tends to decrease the demand for physician visits, implying that physician visits and hospital care are substitutes. Controlling for health status, age decreases the probability of use but is unrelated to demand conditional on use. Females demand more medical care than males. No strong relationships are found between the demand for medical care and income or health beliefs.

As the study for medical care demand goes deeper and broader, Kenkel (1994) investigates the health capital investment motives. First, the analysis indicates that annual use of preventive services decreases with age, which is consistent with individuals rationally reducing their use of preventive care as the pay-off period to the investment shortens over the lifecycle. Second, schooling is found to be an important determinant of demand, with the more educated likely to use additional preventive services. Neither lifecycle nor schooling effects are consistently found in studies of the demand for curative care.

Among the vaccine studies of the 1990s, Kenyon *et al*. (1998) investigates children's vaccine rate in different geographic regions. Using a multistage cluster sampling method in a household survey, they compare vaccination coverage among children 19 to 35 months of age from public housing developments, where a free vaccine outreach program was in place, with children residing elsewhere in the city. Based on their comparison, they found that coverage was significantly lower among children residing in public housing compared with those residing in high-risk strata and low-risk strata. From what they found out, they concluded that African-American children throughout Chicago, particularly in public housing, remain at increased risk for vaccine-preventable diseases and should be targeted further for vaccination services. Their findings indicate the importance to put more attention on the relationship between the geographic factors and vaccine demand. Later vaccine demand studies also demonstrate that cluster surveys are useful for monitoring vaccination coverage in high-risk urban settings.

In the 21st century, immunization rates rose steadily. As a result, some of the diseases that once raged across the country, such as measles, rubella, and polio, no longer occur naturally in the USA. Nearly 81% of American babies get all their recommended vaccinations before age three.

Since achieving and sustaining high immunization rate among U.S. children is an important public health goal that has always been vigorously pursued, extensive research on the factors that contribute to underimmunization has let to a variety of interventions. However, little attention has been paid to extraimmunization, i.e., vaccine doses given in excess of the recommended schedule. Feikema *et al*. (2000) provided quantitative estimates of the cost of extraimmunization in U.S children and identified its associated factors. Their investigation indicated that extraimmunization can be costly. The challenge is to reduce extraimmunization without interfering with more important efforts to combat underimmunization. They were among the earliest resources who realized the vaccine problem.

Davis *et al*. (2001) conducted a national survey regarding childhood vaccine risk/benefit communization in private practice office settings. Their survey indicated a mismatch between the legal mandate for Vaccine Information Statement distribution and the actual practice in private office settings. The majority of providers reported discussing some aspect of vaccine communication but 40% indicated that they did not mention risks. Legal and professional guidelines for appropriate content and delivery of vaccine communication need to be clarified and to be made easily accessible for busy private practitioners. Efforts to improve risk/ benefit communication in private practice should take into consideration the limited time available in an office well-infant visit and should be aimed at both the nurse and physician.

In the Kutty (1999) model, elderly functionality is produced with the direct inputs of assistive devices, personal assistance, and nutritional intake. Education, endowment variables, and health conditions determine the production function environment. His results suggest that reverses in functionality caused by age and health conditions can be partially compensated for by the use of assistive devices, secure nutritional intake, and moderate alcohol consumption, even though non-inputs like chronic health conditions, age, sex and genetic endowment very strongly determine the level of functionality. His study is the extension of Kenkel's medical demand model.

Stokley *et al*. (2001)'s estimated the vaccination coverage levels of children living in rural areas and identified statistically significant differences in coverage between children living in rural areas and their suburban and urban counterparts. Their empirical method was similar to the Kenyon *et al*. (1998) analysis toward children's vaccine rate in different geographic regions. Their conclusions were also consistent with Kenyon *et al*.: children living in rural areas are just as likely to receive the basic 4:3:1:3 vaccination series as their suburban and urban counterparts. Uptake of the varicella vaccine appears to be slower in rural areas than urban areas. Why vaccine update is generally slower in rural areas than urban areas? Further studies are recommended to identify the risk factors for not receiving the vaccine in rural areas.

By age two a child who is up-to-date for immunizations will have received up to 19 shots delivered over eight visits at a market cost of $525 dollars for the vaccines alone, a far more expensive and demanding regimen that the eight shots received in 1987. In recognition of the potential importance of health insurance to immunization coverage rates, Joyce and Racine (2003) use data from the National Immunization Survey for the years 1995 to 2001 to test whether the State Children's Health Insurance Program (SCHIP) was associated with differential changes among poor and near-poor children relative to their non-poor counterparts in either age-appropriate immunization rates or in the proportion of vaccines delivered by private providers. They indicate the probability that a child was up-to-date for the varicella vaccine increased between 7 and 16% more among poor and near-poor relative to non-poor children after implementation of SCHIP. The increase was greater among children from urban areas, among Hispanics, and among those from states with the highest rates of uninsured children prior to SCHIP than among children nationally. They found small to inconsequential changes for other vaccines. They also found that the probability that a poor or near-poor child obtained all vaccines at a private provider fell relative to the same probability among non-poor children over the study period. SCHIP appears to have affected the uptake of a recently introduced vaccine, which suggests that insurance coverage may be important for the rapid adoption of the latest and increasingly more expensive agents such as the pneumococcal conjugate vaccine. The Joyce and Racine study and finding are consistent with previous conclusion about the relationship between the geographic factor and the up-to-date vaccine status, however, they indicate that household income also affects children's vaccine demand, which is contrary to the Kenkel's result "No strong relationships are found between the demand for medical care and income or health beliefs."

Kremer and Snyder (2003) implied that a pharmaceutical manufacturer would have the same incentive to develop either vaccines or drug treatments given they were found to yield the same revenue for a pharmaceutical manufacturer. By embedding an economic model within a standard dynamic epidemiological model, it was shown that the firm can make arbitrarily higher revenue in percentage terms with drug treatments than with vaccines.

Losasso and Buchmueller (2004) wrote a similar article as Joyce and Racine (2003). They used the type of health insurance under which the child is covered as a dependent variable and came to a conclusion that the enactment of SCHIP led to a small but statistically significant reduction in the rate of uninsurance for children. Again, he emphasized that the insurance status will matter on children's vaccination behavior.

Pediatric immunization rates have increased in the U.S. since 1990. Nevertheless, national survey data indicate that up to one third of two-year-old children in some states and urban areas lack at least one recommended dose of diphtheria–tetanus–pertussis (DTP)-, polio-, or measles-containing vaccines. Immunization has become a key measure of preventive pediatric health care in the U.S. To achieve and maintain the national immunization goal that 90% of children receive all recommended immunizations by two years of age, Morrow *et al*. examined access to pediatric immunization services and health system factors associated with underimmunization in a representative sample of children at 12 and 24 months of age. They defined UTD (up-to-date) as an indicator variable for whether a child is up-to-date in vaccine status. It is an indicator for whether the child has more than four doses of DTaP/DTP/DT (The DTaP/DTP/DT vaccine protects your child against three diseases: diphtheria, tetanus, and pertussis (whooping cough))., three doses of polio, one dose of meningococcal conjugate vaccine (MCV), three does of Haemophilus influenza B (HIB), three doses of hepatitis B and one dose of varicella after he/she is one year old. They found that household risk factors for children not being UTD at 12 and 24 months included having a greater number of children, single parenthood, lack of education beyond high school, teenage mother, African-American ethnicity, and not finding the child's immunization record at home.

Browngoehl *et al*. evaluate the impact of an immunization outreach program on immunization rates. Their empirical results indicate that members with home visits have significantly higher completed immunization rates than do other members. Their finding provided evidence supporting a correlation between comprehensive strategies (computerized tracking, member and provider education and incentives, and home visiting) and increased immunization rates. Those individuals who received home visits were more likely to complete an immunization series by 35 months of age than those who did not.

Based on the literature, a high insurance coverage, a high level of parents' education, an abundant knowledge about vaccines, and high household income are possible factors that increases children's vaccine demand.

On the other hand, low insurance coverage, poorly-education level of parents, little knowledge about immunization, and low household's income are possible factors that may cause the household's demand for children's vaccine to decrease.

Furthermore, age of parents, race/ethnicity of the parents, marriage status of mother, whether the child was the first to be born, geographic sectors, physician's visit, drug treatment, and health program attendance are factors that may cause the vaccine demand to change.

CHAPTER 3

DATA AND THEORETICAL MODEL

Description of Data

The analysis of this thesis is based on data from the National Immunization Survey (NIS) 2003 Public-Use Data File. The NIS is sponsored by the National Immunization Program (NIP) and conducted jointly by NIP and the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC). The NIS uses two phases of data collection to obtain vaccination information for a large national probability sample of young children: a random-digit-dialing (RDD) survey designed to identify households with children 19 to 35 months of age, followed by the Provider Record Check Study (PRC), which obtains provider-reported vaccination histories for these children.

The NIS is a national probability sample of children ages 19 to 35 months within the U.S. There are 78 Immunization Action Plan (IAP) areas representing the 50 states and 28 metropolitan areas in the NIS. Each IAP is a stratum and households are drawn randomly within each stratum. Approximately 420 households were surveyed in each of the 78 strata in 2003 for a total of approximately 34,000 households. The survey is by phone and uses a RDD design to identify households with children of the appropriate age. The person most knowledgeable about the child's immunizations is asked to be the respondent. Each respondent is asked to locate an immunization card if available. In addition to information on the number and date of vaccines, respondents are asked about maternal schooling, family income, marital status, and other sociodemographic information. Respondents are then asked permission to contact their immunization providers. The second survey within the NIS is the PRC. Providers are mailed a survey in which they are asked to furnish the child's vaccination history. The initial mailing is followed up with reminders and telephone calls.

The provider data are generally considered the most reliable. However, complete provider data are obtained for only 65% of the 34,000 households surveyed. Those with complete provider data are more likely to be white, better educated, and have higher incomes than households without provider data. Aware of the potential problems associated with selective reporting, administrators of the NIS use propensity scores within each stratum to adjust sampling weights for households with non-provider data. The major concern of our analysis is the probability of being up-to-date for various vaccines. All respondents with complete provider data are used and those that do not know and refused to answer are eliminated. Thus, if a child is reported up-to-date for a vaccine based on the provider survey, then it is assumed the child is up-to-date for the vaccine. The NIS has three categories for the child's age: 19 to 23 months, 24 to 29 months, and 30 to 35 months. Most vaccines are initiated in infancy. Measles, mumps and rubella (MMR) and varicella vaccines begin at 12 months of age.

When determining coverage, up-to-date (UTD) status was used rather than number of doses because the doses required for being UTD varies depending on timing of vaccinations, area requirements regarding number of doses, and brand of vaccines. Consequently, a household composite variable UTD is used as the dependent variable. UTD is an indicator variable for whether a child is up-to-date in vaccine status. It is an indicator for whether the child has at least four doses of DTaP/DTP/DT, three doses of polio vaccine, one dose of MCV, three doses of Haemophilus influenza B (HIB), three doses of hepatitis B and one dose of varicella after he/she is one year old. UTD is a dummy variable, its value is zero for those who are not up-to-date and one for those up-to-date. The primary concern of a study by Morrow *et al*. (1998)'s study was also UTD immunization status of study children for the combination of DTP, polio, and measles–mumps–rubella (MMR) vaccines.

Key demographic and socioeconomic variables in 2003 NIS data file are used based on the literature references. These variables are hypothesized to be factors that affect households' compliance in childhood immunization. Table 3.1 lists the variables used, their meanings and categories in the data set. Key demographic variables include race/ethnicity category of the child (Raceethk), number of children in the household (Childnum), age category of the child (Agegrp), gender of child (Sex), age category of the mother (M_agegrp), marital status category of the mother (Marital), and firstborn status of the child (Frstbrn). Davis *et al*. (2001) have similar demographic variables in their analysis investigating the risk and benefits of childhood immunization. Stokley *et al*. (2001) also employ similar explanatory variables including household size and firstborn status of child in their research regarding the vaccine status for children from different regions.

Key socioeconomic variables include education category of mother (Mumeduc), and the income-to-poverty ratio (Incporat). Dubay and Kenny (2003) included these two socioeconomic variables in their research regarding the effect of Medicaid program on childhood immunization. Feikema *et al*. (2000) investigated the cost of extra immunization also investigated these two socioeconomic variables.

The federal program for women, infants, and children (WIC) variable is also considered, which describes whether the child ever participated in the WIC program. WIC, a Federal grant program for which Congress authorizes a specific amount of funds each year, serves 45% of all

Table 3.1. Dependent Variable and Explanatory Variables [a]

Variable Names	Description	Categories	
Dependent Variable			
UTD	Up-to-date flag for provider 4:3:1:3:3:1 (4 DTaP/DTP/DT, 3 polio, 1 MCV, 3 Hib, 3 hepatitis B and 1 varicella)	0	Not up-to-date
		1	Up-to-date
Independent Variable			
Agegrp	Age category of child	1	19-23 Months
		2	24-29 Months
		3	30-35 Months
Childnum	Number of children in household	1	Child
		2	2-3 Children
		3	4+ Children
Wic	Children ever received WIC benefits	0	No
		1	Yes
Mumeduc	Education of mother categories	0	No College
		1	College
Frstbrn	First born status of child	0	No
		1	Yes
Incporat	Income to poverty ratio	0.5	Minimum value
		3.0	Maximum value
Marital	Marital status of mother categories	0	Not Married
		1	Married
M_agegrp	Age of mother categories	1	<=19
		2	20-29
		3	30+
Raceethk	Race/Ethnicity of child	0	Non-Hispanic white only
		1	Others
Sex	Gender of child	0	Male
		1	Female
Provnum	Number of vaccination providers identifies by respondent	0	0
		1	1
		2	2
		3	3+
Compcare	Children's provider offers comprehensive child care	1	All providers
		2	Some but not all providers
		3	No provider
Medicaid	Medicaid expansion	0	No
		1	Yes

Table 3.1. (Continued)

Variable Names	Description	Categories	
Separate	Separate new insurance program	0 1	No Yes
Wideflu3yr	Number of times when influenza activity is widespread from 2000-2002	0 20	Minimum value Maximum value
Wideflu00	Number of times when influenza activity is widespread in 2000	0 8	Minimum value Maximum value
Wideflu01	Number of times when influenza activity is widespread in 2001	0 10	Minimum value Maximum value
Wideflu02	Number of times when influenza activity is widespread in 2002	0 9	Minimum value Maximum value

[a] Also includes the 50 state effects dummy variables.

infants born in the United States. Joyce and Racine (2003) indicate such a program will have a major impact on narrowing the immunization coverage rate gaps among poor and near-poor children relative to their non-poor counterparts. Morrow *et al*. (1998) discovered enrollment in WIC is associated with significantly increased immunization rates.

The immunization history of each child is determined mainly from parent and/or provider records. The combination of parent and provider records, when available, is used to provide a complete immunization history for each child. Thus, some typical medical care variables are also used in the model: number of vaccination providers identified by respondent (Provnum) and child's providers offer comprehensive child care (Compcare). Based on previous research, vaccination is one of the most important choices if child's providers offer comprehensive child care. Joyce and Racine (2003) used an indicator to identify whether the child received the vaccine in a comprehensive care setting. They indicate a direct link between the comprehensive child care and the up-to-date coverage rate.

Dummy variable on Medicaid (Medicaid) and separate new insurance program (Separate) are included in the model. Medicaid is a program that pays for medical assistance for certain individuals and families with low incomes and resources. This program became law in 1965 and is jointly funded by the Federal and State governments (including the District of Columbia and the Territories) to assist states in providing medical long-term care assistance to people who meet certain eligibility criteria. Medicaid is the largest source of funding for medical and health-related services for people with limited income. A number of recent studies found high up-to-date childhood immunization rates for individuals who became eligible for public insurance through the Medicaid expansions of the late 1980s and early 1990s. Dubay and Kenny (2003) concluded that expanding public health insurance coverage to parents has benefits to children in the form of increased participation in Medicaid. However, some recent studies also suggest that the Medicaid expansions contributed to a decline in private insurance, though the estimated magnitude of this effect varies considerably. LoSasso (2004) found that Medicaid has a large crowded-out effect on several new insurance programs. Thus, whether the crowd-out effect Medicaid has on insurance program is more or less of a problem in the case of childhood vaccination is an empirical question.

Variables that measure how often influenza occurs in each state are developed. The variables are based on the data from the Weekly Surveillance Reports of Influenza Branch at CDC from 2000-2002. State health departments report the estimated level of influenza activity in their states each week. States report influenza activity as no activity, sporadic, local, regional, or widespread. Widespread flu activity is said to occur when there are outbreaks of influenza-like illness or culture-confirmed influenza in counties having a combined population of at least 50% of the state's population. From the CDC Flu activity webpage (http://www.cdc.gov/flu/weekly/fluactivity.htm) maps are obtained which illustrate influenza activity as assessed by State and Territorial epidemiologists. A sample map for the week ending Feb 12, 2000 is provided in Figure 3.1. From the map, four states are colored red (widespread) for the week, indicating influenza activity was reported as widespread. These four states are: Arizona, New York, Pennsylvania, and Tennessee. Based on these maps, a spreadsheet is created indicating influenza activity by state. A five is coded for those states where influenza activity was widespread. Counting the number of fives for each year created four variables: number of fives for three years (2000-2002) in a row (Wideflu3yr) and number of fives for three separate years (Wideflu00, Wideflu01 and Wideflu02). Healthy children aged 6 to 23 months are eligible to receive influenza vaccine, since they are at increased risk of hospitalization related to

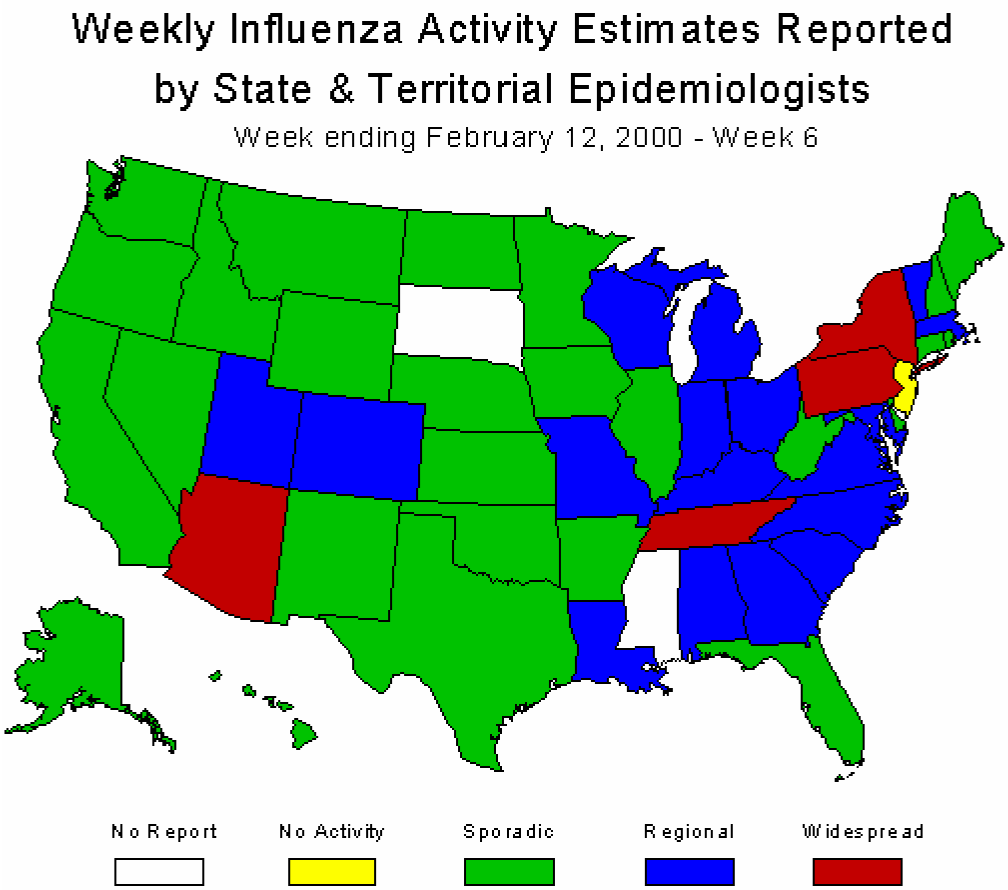

Figure 3.1. Weekly Influenza Activities

influenza compared to older children and young adults. The households of these children who experienced influenza before tend to keep their children more up-to-date in many other vaccines besides influenza vaccine (Richards and Stemnock (2005)). These children are potentially at much higher risk than their counterparts. Mullahy (1998) used a similar variable, the extent to which the individual is exposed to flu virus in his research addressing people's willingness to get immunized against the flu.

Philipson (1995) investigated the degree to which the occurrence of vaccine-preventable diseases affects vaccination efforts against such diseases and found out that the occurrence of vaccine-preventable diseases can stimulate vaccine demand. In particular, Philipson's paper documents and discusses the importance of what is referred to as a prevalence elastic for preventions such as vaccines, meaning that the demand for vaccination rises with the prevalence of the disease. Philipson also stressed that the demand for vaccination is represented by the hazard rate from susceptibility immunity. Hazard rate is denoted by $h(Z_t, p)$, where Z_t is a future prevalence path and p is the price of vaccination. Hazard rate is based on the fraction of susceptible individuals that vaccinate at any given time. The demand for vaccination is represented by the hazard rate as an individual vaccinate and move from being susceptible to the disease and develop immunity to the condition. Philipson classified agents into three categories: susceptible, infected, and recovered through immunity, denoted S, I, and R, respectively. He denoted the transition of a particular disease between two health states by $\lambda_{\alpha\alpha'}$, the fraction of susceptible that become infected (λ_{si}) equals to the product of the probability of transmission between susceptible and infected (β) times the proportion of the total population that is infected (I_t). Then according to the law of large numbers, Philipson developed his deterministic system represented by equation 3.1.

$$\begin{cases} dS_t / dt = \lambda_{os} - \beta I_t S_t - \left(\lambda_{sd} + h(Z_t, p)\right) S_t \\ dI_t / dt = \beta I_t S_t - \left(\lambda_{ir} + \lambda_{id}\right) I_t \\ dR_t / dt = \lambda_{ir} I_t + h(Z_t, p) S_t - \lambda_{rd} R_t \end{cases} \qquad [3.1]$$

Philipson assumes that the hazard rate increases as the number of infected people rises, where I represent the level of infection. The hazard rate (demand for vaccinations) decreases as the price of the vaccination rises. These assumptions imply that $h_I \geq 0$ and $h_p \leq 0$. The deterministic system implies that the change in the stock of infected individuals is due to the entry of new infections, $\lambda_{si} S_t$, exits due to natural immunity at rate λ_{ir} and exits due to nondisease-related deaths at rate λ_{id}. Also, the change in the stock of immune individuals is due to the entry of individuals recovering from infections, $\lambda_{ir} I$, the entry of vaccinated susceptible, hS_t, and exits due to nondisease-related deaths, $\lambda_{rd} R$. The empirical results are presented on the prevalence elasticity in Chapter 4 compared to Philipson's findings.

Stokley (2001) identified statistically significant differences in vaccine coverage between children living in rural areas and their suburban and urban counterparts. This indicates the possibility of different childhood vaccine compliance in different states. Thus, 51 state effects dummy variables (50 states plus District of Columbia) are created and each of these variables works as a state indicator.

Table 3.2 lists the descriptive statistics for the variables. From the table, it is observed that among the 19,565 households, there are about 72.4% households that are up-to-date in childhood immunization. The average age of the child is above 29 months. Each family usually has more than one child. Also, about 57.8% of the mothers of the household have never attended college. Since there is generally more than one child in the households, the rate of firstborn among these children is as low as 39.7%. From the income to poverty ratio, it is observed that

Table 3.2. Descriptive Statistics for the Variables [a]

Variable [b]	Mean	Standard Deviation	Minimum	Maximum
UTD	0.724	0.447	0	1
Agegrp	2.052	0.803	1	3
Childnum	1.854	0.604	1	3
Wic	0.468	0.499	0	1
Mumeduc	0.422	0.494	0	1
Frstbrn	0.397	0.489	0	1
Incporat	2.049	0.961	0.5	3
Marital	0.812	0.391	0	1
M_agegrp	2.549	0.537	1	3
raceethk	0.418	0.493	0	1
Sex	0.483	0.500	0	1
Provnum	1.382	0.594	1	3
Compcare	1.227	0.582	1	3
Medicaid	0.340	0.474	0	1
Separate	0.253	0.435	0	1
Wideflu3yr	6.201	6.017	0	20
Wideflu02	2.227	3.144	0	9
Wideflu01	1.110	2.066	0	10
Wideflu00	2.864	2.311	0	8

[a] There are 19,565 observations.

[b] Refer to table 3.1 for a description of the variables.

most American families have a very stable income. There is little difference between the mean value and the maximum value. Most mothers among these households are above thirty years old. Most households (around 58%) are non-Hispanic white among the 19,565 observations and there are more females (51%) than males. On average, there are 1.382 vaccination providers for a household and some of them always offer comprehensive child care. Approximately 34% of the households participate or have participated in Medicaid program and about 25.3% ever received separate insurance program.

Approximately 35% of the children within the households surveyed are in the 24-29 months category, also 35% of children are in the 30-35 months category but only 30% of children are in the 19-23 months category. This indicates the population of children in the data set is generally uniformly-distributed. Furthermore, females and males hold almost the same proportion in our dataset. Two-thirds of American households have more than one child. Among all the respondents, 60% are non-Hispanic white households and 40% are not. There is a substantial difference in sample population between whites and minorities. In the mother categories, 42% have never received a college education, 58% have ever attended college. Most households (68%) have only one vaccination provider, 27% have two and only 5% have more than three. All providers can offer child comprehensive child care among 85% of the households. More than 70% of households are up-to-date, indicating that most households realize the importance of vaccination.

In table 3.3, a number of cross tabs are listed for Agegrp, Childnum, and Compcare to the UTD outcome. From table 3.3, childhood immunization compliance increases as the age of the children increases. The oldest age group of children who are between 30 to 35 months old and the 2nd oldest age group of children, who are between 24 to 29 months old have the highest

Table 3.3. Cross-Tabs Between Up-to-date Vaccine Status and Several Variables [a]

	Up-to-date in childhood vaccination (UTD)		
Variable name[b]	**Not up-to-date**	**Up-to-date**	**Total**
Ageprp			
19-23 months	10%	20%	30%
24-29 months	9	26	35
30-35 months	9	26	35
Total	28	72	100
Childnum			
1	6	20	27
2-3	17	45	61
3+	4	8	12
Total	28	72	100
Compcare			
All providers	23	63	86
Some providers	2	5	6
No providers	3	5	8
Total	28	72	100

[a] Variables are chosen based on the results from the logit regression.

[b] Agegrp is the age category of child. Childnum is the number of children in the household. Compcare is children's provider offers comprehensive child care.

up-to-date immunization rates of 26%. The youngest age group of children who are between 19 to 23 months old has the lowest up-to-date immunization rate of 20%. At birth, infants have immunity to certain diseases because antibodies have passed through the placenta from the mother to the unborn child. After birth, the breastfed baby gets the continued benefits of additional antibodies in breast milk. But in both cases, the immunity is only temporary. Immunization (vaccination) is a way of creating immunity to certain diseases - by using small amounts of a killed or weakened microorganism that causes the particular disease. Research results regarding the relationship between the age of child and the vaccination rate is consistent with Browngoehl *et al*. (1997)'s research outcome. They indicate that as children age from 20 months to 35 months, their parents tend to provide them with vaccines. Parents are less and less concerned about serious side effects or reactions that their children may get.

A conclusion can also be drawn on the correlation between the number of children in the household and vaccination compliance. As indicated in table 3.3, those households that have three children or more have the lowest compliance vaccination rate of 8%. Note that the household whose children are in the second children's number category have a higher up-to-date vaccination rate. It is not true because the second children's number category has a bigger population than the first children's number category. Its up-to-date rate is the summation of the vaccine up-to-date rate of households that have two children and the vaccine up-to-date rate of households having three children. It can also be inferred from the outcome of logit model that households that have two or three children must have a lower up-to-date childhood immunization rate than the households that have only one child. This also makes sense because a household will reschedule its budget given an extra child.

Using a similar method, a conclusion can be drawn on the relationship between the up-to-date vaccine status and characteristics of provider variables. Joyce and Racine (2003) found the relationship between the probability that a child was up-to-date for the varicella vaccine and the number of vaccination providers identified for the child. Similarly, the analysis examines whether children receive vaccines from a single provider, a single private provider, or one that provides comprehensive pediatric services. The findings are that the probability of a poor or near-poor child obtained all vaccines at a private provider fell relative to the same probability among non-poor children over the study period. From table 3.3, if all the providers of the child can provide comprehensive pediatric services, up-to-date childhood immunization rate is very high (63%). This is consistent with Joyce and Racine's household survey that asked each respondent how many vaccine providers they used. He used those who saw only one provider as a positive indicator of continuity.

Theoretical Model

A household's decision to be in compliance with a child's immunization is an example of the demand for preventive care under uncertainty. The household does not know for certain the risks and effectiveness of immunization. The decision is then a discrete choice in which the expected net welfare from being in compliance is compared with the expected net welfare of noncompliance.

Figure 3.2 illustrates the consequences of the immunization decision. The household has the choice of having its child being in compliance with immunizations or not. Given this choice there is a joint probability of the child becoming ill or well depending on if the child is in compliance or not. The household will then make the decision to be in compliance depending on the probabilities it assigns to the possible outcomes and magnitude of the associated

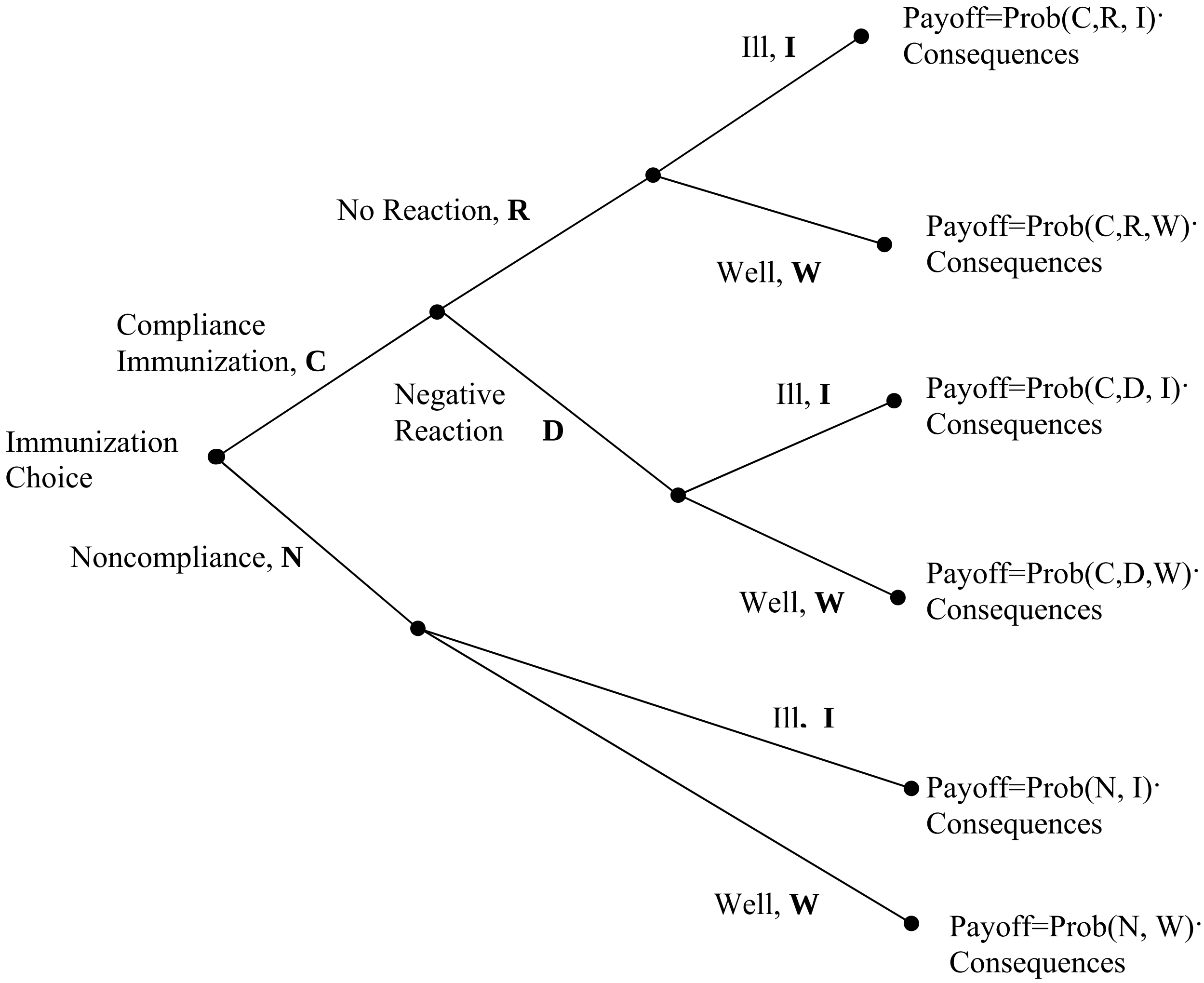

Figure 3.2. Decision Tree

consequences. The household determines these probabilities and consequences (outcomes) based on a set of exogenous variables, X, including all the explanatory variables but mother's education (Mumeduc). In particular, it is hypothesized mother's education, Mumeduc, has a direct effect on these outcomes, called the endowment effect, and an indirect effect, called the coefficient effect. Increased education can directly influence the outcomes but also indirectly influence the outcomes by changing the values of X. The household will then make a choice y_i (i = 1 if the choice is to be in compliance and 0 if to be in noncompliance) if it yields the highest expected utility

$$EU\left[y_i \mid Mumeduc, X(Mumeduc)\right] > EU\left[y_j \mid Mumeduc, X(Mumeduc)\right], \qquad [3.2]$$

where E is the expectations operator and U represents Utility. The uncertainty is a consequence of the probabilistic nature of immunization outcomes.

The analysis can then be cast in a discrete choice econometric model by appending additive stochastic elements to the utility functions

$$W\left[y_i \mid Mumeduc, X(Mumeduc)\right] = EU\left[y_j \mid Mumeduc, X(Mumeduc)\right] + \varepsilon_i . \qquad [3.3]$$

This theoretical specification decomposes the effect education has on a household's immunization compliance level. It provides the underlying theoretical foundation for measuring the relative magnitude of the joint effect of education on compliance, the coefficient plus endowment effect, and the share that each contributes. As discussed in the following section, the logit model is employed for estimating this discrete choice model and based on the results, the Oaxaca-Blinder decomposition approach is used to determine the relative share of education's effect on compliance due to the coefficient and endowment effects.

Model and Methodology

Many variables in the social sciences follow binomial distributions. Binary outcomes such as voted or not voted, dead or living, agreed or disagreed, migrated or not migrated, and the general occurrence or nonoccurrence of an event all generate binomial distributions. Logit and probit models are often used to study these events.

The dependent variable used for our model, UTD, is also a dummy variable. Its value is either 0 (for not up-to-date) or 1 (for up-to-date). Despite similarity of logit and probit, there are two practical advantages of the logit model. The first one is simplicity. The equation of the logistic cumulative distribution function (CDF) is very simple, while the normal CDF involves an unevaluated integral. This difference is trivial for dichotomous data, but for polytomous data, which requires the multivariate logistic or normal distribution, the disadvantage of the probit model is much more acute. The second one is interpretability: The inverse linearizing transformation for the logit model is directly interpretable as a log-odds, while the inverse transformation for the probit model does not have a direct interpretation. Considering the advantages in simplicity and interpretability, the logit model is used for analysis.

The logit model starts from the assumption that the underlying response variable y^* is defined by the regression relationship

$$y^* = \sum_{k=1}^{K} \beta_k x_k + \varepsilon. \qquad [3.4]$$

In practice, y^* is unobserved, and ε is symmetrically distributed with zero mean and has its CDF defined as $F(\varepsilon)$. What is observed is a dummy variable y, $y = 1$ if $y^* > 0$; $y = 0$ otherwise. In the model formulated by the above equation, the summed term of $\beta's$ and $x's$ is not $E(y|x_1,\ldots,x_k)$ as in the linear case, but $E(y^*|x_1,\ldots,x_k)$. From these relations

$$\Pr(y=1) = \Pr\left(\sum_{k=1}^{K} \beta_k x_k + \varepsilon > 0\right)$$
$$= \Pr\left(\varepsilon > -\sum_{k=1}^{K} \beta_k x_k\right) = 1 - F\left(-\sum_{k=1}^{K} \beta_k x_k\right). \qquad [3.5]$$

where F is a CDF of ε. Denote η, which is defined in the following equation 3.4 (μ is the expected probability), the generalized linear predictor, as the systematic component in y^*, and ε as the random component in y^*. The functional form of F depends on the distribution of, or rather, the assumption made about the distribution of ε. The distribution of ε determines the link function of a generalized linear model, another way to represent F. Assuming the random component of the response in the data follows a binomial distribution and assuming the logistic distribution for ε, the logit model can be applied to the data. The link function then becomes the logit:

$$\eta = \log[\mu/(1-\mu)]. \qquad [3.6]$$

Applying this link function specifies a logit model that takes two forms. When expressed in logit form, the model is specified as:

$$\log\left[\frac{P(y=1)}{1-P(y=1)}\right] = \sum_{k=1}^{K} \beta_k x_k. \qquad [3.7]$$

For an event, it is

$$\Pr(y=1) = 1 - L\left(-\sum_{k=1}^{K} \beta_k x_k\right) = L\left(\sum_{k=1}^{K} \beta_k x_k\right) = \frac{e^{\sum_{k=1}^{K} \beta_k x_k}}{1+e^{\sum_{k=1}^{K} \beta_k x_k}}. \qquad [3.8]$$

For a nonevent, the probability is just 1 minus the event probability or:

$$\Pr(y=0) = L\left(-\sum_{k=1}^{K} \beta_k x_k\right) = \frac{e^{-\sum_{k=1}^{K} \beta_k x_k}}{1+e^{-\sum_{k=1}^{K} \beta_k x_k}} = \frac{1}{1+e^{\sum_{k=1}^{K} \beta_k x_k}}. \qquad [3.9]$$

The two forms of the logit model are responsible for the different names given to this type of modeling. The model expressed as [3.7] leads to “logit models” because of the logit term, while that expressed as [3.8] leads to “logistic regression” because of the cumulative logistic distribution function. In the thesis, the term “logit models” refers to both forms.

Given the logit model, the focus turns to selecting the explanatory variables. Besides the key variables in the NIS data set, the variable Medicaid and Separate are employed from the research by LoSasso and Buchmueller investigating the impact of the State Children’s Health Insurance Program (SCHIP) on health insurance coverage. The SCHIP was signed into law as part of Title XXI of the 1997 Balanced Budget Act. The goal of the legislation was to increase the insurance coverage of children in the United States by extending eligibility for public insurance to children in families earning too much to qualify for Medicaid yet earning too little to afford private health insurance. Touted as the largest expansion in health insurance since the enactment of Medicaid in 1965, the SCHIP legislation apportioned more than $40 billion in federal matching funds over ten years beginning in fiscal year 1998. States are very flexible in using these funds. They are allowed to use these funds to expand Medicaid eligibility, to develop new insurance programs, and increase outreach for children already eligible for public coverage.

As referred by LoSasso and Buchmueller, states had three broad options for implementing SCHIP. They could expand their Medicaid programs by either increasing income eligibility thresholds or extending coverage to age groups that were not eligible for Medicaid previously, create a new separate health insurance program for children, or do both. As of March 2001, 19 states expanded Medicaid, 15 states created a separate SCHIP program, and 17 states implemented a combination program. The table “Summary of SCHIP expansions, by state for the years 1996 and 2000” in LoSasso and Buchmueller’s article summarizes the timing of SCHIP implementation and how the program's effect on public insurance eligibility varied across states and age groups. Most states (34) enacted their program in 1998. Eleven states did so in 1997 and the remaining six states began in 1999 or 2000. States that implemented both Medicaid expansions and a separate SCHIP program were able to start each component at different times, usually expanding Medicaid eligibility first. For analysis, the expansion type (expansion) variable is, one for Medicaid expansion only, two for separate new insurance program and three for combination program. This variable was merged into the NIS data set with the recoded table in LoSasso and Buchmueller. This is the basis for creating two dummy variables, Medicaid expansion (Medicaid) and separate new insurance program (Separate). Medicaid equals to one when expansion equals to one. Separate equals to one when expansion equals to two.

LoSasso and Buchmueller also have an impact on building the regression model. Their baseline model:

$$COVERAGE_{ci} = \alpha_c PublicEligibility_i + \beta_c X_i + \gamma_c STATE_i + \theta_c TIME_i + \varepsilon_{ci}, \qquad [3.10]$$

where the dependent variable $COVERAGE_{ci}$ represents the type of health insurance under which the child is covered (c = public, private, or uninsured). Public Eligibility is an indicator for public insurance eligibility. The vector X contains demographic variables, such as the child's gender, race, age, the number of people in the family. They also include a full set of year dummies to account for national trends in health insurance coverage and state dummies to capture long-standing differences across states in economic conditions, health care market characteristics, and normal UTD differences. Joyce and Racine (2003) used state and year fixed effects in their research on the association between the state children's health insurance program and immunization coverage and delivery.

The starting point for the econometric analysis is the following regression model:

$$UTD = \alpha X_i + \beta STATE_i + \gamma Wideflu02_i + \lambda Wideflu01_i + \theta Wideflu00_i + \varepsilon_i . \qquad [3.11]$$

where the dependent variable UTD_i represents the indicator for whether the child is up-to-date in vaccine status. The vector X contains household variables, indicators related to the mother, medical care variables, and state Medicaid program variables, including age category of the child (Agegrp), whether child ever participated in the WIC program (WIC), number of children in the household (Childnum), firstborn status of the child (Frstbrn), the income-to-poverty ratio (Incporat), race/ethnicity category of the child (Raceethk), age category of the mother (M_agegrp), marital status category of the mother (Marital), education category of mother (Mumeduc), number of vaccine providers in the household (Provnum), child's providers offer comprehensive child care (Compcare), Medicaid program (Medicaid) and separate insurance program (Separate). STATE is the state effect dummy variable used to study the differences across the states in up-to-date vaccine status. Wideflu02, Wideflu01 and Wideflu00 are the number of times when influenza activity is widespread in 2000, 2001 and 2002 respectively. Wideflu3yr is the number of times when influenza activity is widespread from 2000 to 2002.

The model that contains Wideflu3yr, which is the number of times when influenza activity is widespread from 2000 to 2002 can be described as the following

$$UTD_i = \alpha X_i + \beta STATE_i + \gamma Wideflu3yr_i + \varepsilon_i , \qquad [3.12]$$

The logit model in STATA is used to estimate equations 3.9 and 3.10. The method of maximum likelihood is used by STATA to estimate the parameters in general logit models. Given the parameters, the logit regression model specifies how to calculate the probability that a specific outcome will occur (e.g. $y_1 = 1, y_2 = 0, \ldots, y_n = 0$, where n is the number of observations). Given n independent responses, with π_i the mean response for observation i then

$$\Pr\left(Y_1 = y_1, \ldots, Y_n = y_n\right) = \prod_{i=1}^{n} \pi_i^{y_i} (1-\pi)^{1-y_i}. \quad [3.13]$$

The likelihood function depends on the parameters (β's) through π; the probability of the observed outcome can be computed under various possible choices for the β's.

Oaxaca-Blinder Decomposition

Researchers may be interested in finding out exactly how much a rate difference or change in an explanatory variable is due to compositional difference or change and how much is due to an actual rate difference or change. Decomposition methods allocate explanatory variable rate differences into components of rate and various compositional differences. In her classical research extending standardization to decomposition methods, Kitagawa (1955) demonstrated decomposition when the crude rate difference is confounded by up to three factors, and her two-factor, four-component method has attracted much attention. Later developments lead to refinements based on her decomposition methods.

The decomposition technique is frequently applied in economic problem analysis. The Oaxaca-Blinder decomposition plays a significant role in labor economics. The most popular example is in male-female wage discrimination. The basic idea of the Oaxaca-Blinder decomposition is that differences in the dependent variable, wages, can be explained by two parts. One is what is called "the differences in characteristics", the binary male-female variable, and the other one is what is called "the differences in OLS coefficients". The application potential of this decomposition analysis is not restricted to studying continuous variables such as wage differentials. The variable whose differences in the first moment can be either continuous or discrete.

Oaxaca-Blinder decomposition can show how much of the gap in the dependent variable is due to differing endowments between the two groups such as males/females or educated/uneducated mothers, and how much is due to preferences. For wage differentials this difference in preference is a measure of discrimination. For up-to-date vaccination associated with mother's education, this difference is a mother's or household preference for being up-to-date. As a result, the idea of the Oaxaca-Blinder decomposition has been applied whenever one needs to explain the differences between two comparison groups. To do this researchers usually construct an auxiliary equation based on one group's characteristics and the estimated coefficients of the other group's equation. Detailed decomposition is useful and instructive since it can answer questions like "How much of the differences in vaccination participation rates between different educational groups can be explained by differences in educational attainment (characteristics effect) and how much by differences in behavioral or preference response (coefficients effect)?"

Yun (2003) proposed a simple methodology for decomposing differences in the first moment into characteristics and coefficient effects. This methodology provides a general way to apply the Oaxaca-Blinder decomposition to a non-linear function for both aggregate and detailed decompositions including logit and probit models. Gang *et al*. (2002) applied Yun's methodology and presented the standard errors for decomposition.

Madden (2000) developed techniques to decompose male-female wage differences taking account of preferences in terms of access to the labor market and also selectivity bias. Madden finds considerable evidence of discrimination at point of entry but that discrimination owing to selectivity bias is minimal.

Decomposition is also used in other economics research. Dude (2005) use a modification of the Oaxaca-Blinder decomposition technique for logit models in the article to decompose the shifts in determinants of infant mortality in Kazakhstan into effects due to changes in the relative riskiness of different determinants of infant mortality and to changes in population composition. After examining covariates relating to ethnicity, geographic location, maternal education, household economic status, and characteristics of pregnancies, Dude finds that the most significant changes are driven by the increased likelihood of infant death among non-ethnically Kazakh infants, the decreasing contribution to the birth pool from wealthier households, and the increasing risk to females and infants from multiple births.

Silberman (1992) applied decomposition techniques to environmental economics. By decomposing the difference between users and nonusers of the beach nourishment project, Silberman finds that the greater overall dispersion of the U.S wage distribution reflects considerably more compression at the bottom of the distribution in the other countries, but relatively little difference in the degree of wage inequality at the top.

As early as 1971, Winsborough and Dickinson demonstrated how decomposition works in the three folded division. Based on the regression model

$$Y_j = X_j \beta_j + \varepsilon_j \qquad [3.14]$$

where Y_j is the dependent variable, X_j is the independent variable, β_j is the coefficient of X_j, ε_j is the error term and $E(\varepsilon_j) = 0,\ j \in \{1,2\}$ the mean outcome difference $R = \overline{Y}_1 - \overline{Y}_2 = \overline{X'}_1 \hat{\beta}_1 - \overline{X'}_2 \hat{\beta}_2$ ($\overline{Y}$ is the sample mean of outcome variable (i.e. log wages) and $\overline{X}$ is mean vector of regressors (e.g. education, experience)) can be decomposed as $R = \left(\overline{X}_1 - \overline{X}_2\right)' \hat{\beta}_2 + \overline{X}'_2\left(\hat{\beta}_1 - \hat{\beta}_2\right) + \left(\overline{X}_1 - \overline{X}_2\right)'\left(\hat{\beta}_1 - \hat{\beta}_2\right)$, where $\left(\overline{X}_1 - \overline{X}_2\right)' \hat{\beta}_2$ is differences in endowments, $\overline{X}'_2\left(\hat{\beta}_1 - \hat{\beta}_2\right)$ is differences in coefficients and $\left(\overline{X}_1 - \overline{X}_2\right)'\left(\hat{\beta}_1 - \hat{\beta}_2\right)$ is the interaction.

In 1994, Oaxaca and Ransom simplified the three folded division and developed the two-fold decomposition which is used by researchers currently. The two-fold decomposition can also be expressed as

$$R = \left(\overline{X}_1 - \overline{X}_2\right)' \left[W\hat{\beta}_1 + (I - W)\hat{\beta}_2\right] \text{ (explained part, characteristics effect),}$$
$$+ \left[\overline{X}'_1(I - W) + \overline{X}'_2 W\right]\left(\hat{\beta}_1 - \hat{\beta}_2\right) \text{ (unexplained part, coefficient effect),} \qquad [3.15]$$

where W represents a matrix of relative weights given to the coefficients of the first group (I = identity matrix).

Suppose there are two groups of children with different UTD status, the household of which have different educational level (explanatory variable). The Oaxaca-Blinder decomposition shows: How much of the gap between vaccines status of these two groups is due to difference in endowments (differences in explanatory variables—household's education) and difference in the estimated coefficients (marginal impact of explanatory variables).

Actually, the computation of the decomposition components is standard: Estimate OLS models and insert the coefficients and the means of the regressors into the formulas. However, there currently is no standard method for deriving standard errors for the decomposition components. Few papers applying these methods reports standard errors or confidence intervals. This is problematic because it is hard to evaluate the significance of reported decomposition results without knowing anything about their sampling distribution.

There are several approaches to calculate standard errors in the decomposition. One solution is to use the bootstrap technique. However, bootstrap is slow and it would be desirable to have an asymptotic formulas. Previously proposed procedures produce biased results in most applications because they assume fixed regressors. New unbiased variance estimators for the components of the three-fold and the two-fold decomposition were presented by Jann (2005).

Also, Gang *et al*. (2002) showed the standard errors for decomposition. Using a probit decomposition analysis, they decomposed the difference in the poverty rates between the scheduled castes (or tribes) and non-scheduled households into a part explained by the differences in characteristics and a part explained by the differences in probit coefficients. Standard errors are used to assist detecting the sensitivity of inclusion of the state dummies.

The sampling variance of the mean prediction $\overline{Y} = \overline{X}'\hat{\beta}$ can be estimated if the regressors are fixed, then $\overline{X}$ is constant. Thus:

$$V\left(\overline{X}'\hat{\beta}\right) = \overline{X}'\hat{V}\left(\hat{\beta}\right)\overline{X} \qquad [3.16]$$

In most applications, however, the regressors and therefore $\overline{X}$ are stochastic. Fortunately, $\overline{X}$ and $\hat{\beta}$ are uncorrelated. Thus:

$$\hat{V}\left(\overline{X}'\hat{\beta}\right) = \overline{X}'\hat{V}\left(\hat{\beta}\right)\overline{X} + \hat{\beta}'\hat{V}\left(\overline{X}\right)\hat{\beta} + tr\left(\hat{V}\left(\overline{X}\right)\hat{V}\left(\hat{\beta}\right)\right) \qquad [3.17]$$

where $tr(\cdot)$ is the sum of the characteristic roots of a matrix.

Variance of difference in mean prediction can be estimated as long as the two samples are independent, the variance estimator for the group difference in mean predictions immediately follows as:

$$\begin{aligned}\hat{V}(R) &= \hat{V}\left(\overline{X}_1'\hat{\beta}_1 - \overline{X}_2'\hat{\beta}_2\right) = \hat{V}\left(\overline{X}_1'\hat{\beta}_1\right) + \hat{V}\left(\overline{X}_2'\hat{\beta}_2\right) \\ &= \overline{X}_1'\hat{V}\left(\hat{\beta}_1\right)\overline{X}_1 + \hat{\beta}_1'\hat{V}\left(\overline{X}_1\right)\hat{\beta}_1 + tr\left(\hat{V}\left(\overline{X}_1\right)\hat{V}\left(\hat{\beta}_1\right)\right) \\ &= \overline{X}_2'\hat{V}\left(\hat{\beta}_2\right)\overline{X}_2 + \hat{\beta}_2'\hat{V}\left(\overline{X}_2\right)\hat{\beta}_2 + tr\left(\hat{V}\left(\overline{X}_2\right)\hat{V}\left(\hat{\beta}_2\right)\right)\end{aligned} \qquad [3.18]$$

Similarly, for the three fold decomposition:

$$\hat{V}\left(\left[\overline{X}_1 - \overline{X}_2\right]'\hat{\beta}_2\right) = \left(\overline{X}_1 - \overline{X}_2\right)'\hat{V}\left(\hat{\beta}_2\right)\left(\overline{X}_1 - \overline{X}_2\right) + \hat{\beta}_2'\left[\hat{V}\left(\overline{X}_1\right) + \hat{V}\left(\overline{X}_2\right)\right]\beta_2 + tr(\cdot) \qquad [3.19]$$

$$\hat{V}\left(\overline{X}_2'\left[\hat{\beta}_1 - \hat{\beta}_2\right]\right) = \overline{X}_2'\left[\hat{V}\left(\hat{\beta}_1\right) + \hat{V}\left(\hat{\beta}_2\right)\right]\overline{X}_2 + \left(\hat{\beta}_2 - \hat{\beta}_2\right)'\hat{V}\left(\overline{X}_2\right)\left(\hat{\beta}_2 - \hat{\beta}_2\right) + tr(\cdot) \qquad [3.20]$$

$$\begin{aligned}\hat{V}\left(\left[\overline{X}_1 - \overline{X}_2\right]\left[\hat{\beta}_1 - \hat{\beta}_2\right]\right) &= \left(\overline{X}_1 - \overline{X}_2\right)'\left[\hat{V}\left(\hat{\beta}_1\right) + \hat{V}\left(\hat{\beta}_2\right)\right]\left(\overline{X}_1 - \overline{X}_2\right) \\ &+ \left(\hat{\beta}_2 - \hat{\beta}_2\right)'\left[\hat{V}\left(\overline{X}_1\right) + \hat{V}\left(\overline{X}_2\right)\right]\left(\hat{\beta}_1 - \hat{\beta}_2\right) + tr(\cdot)\end{aligned} \qquad [3.21]$$

Finally, variance of difference in mean prediction is estimated for two-fold decomposition. Assume W is fixed then for the two-fold decomposition,

$$\begin{aligned}\hat{V}(Q) &= tr(\cdot) + \left(\overline{X}_1 - \overline{X}_2\right)'\left[W\hat{V}\left(\hat{\beta}_1\right)W' + (I-W)\hat{V}\left(\hat{\beta}_2\right)(I-W)'\right]\left(\overline{X}_1 - \overline{X}_2\right) \\ &+ \left[W\hat{\beta}_1 + (I-W)\hat{\beta}_2\right]'\left[\hat{V}\left(\overline{X}_1\right) + \hat{V}\left(\overline{X}_2\right)\right]\left[W\hat{\beta}_1 + (I-W)\hat{\beta}_2\right]\end{aligned} \qquad [3.22]$$

$$\begin{aligned}\hat{V}(U) &= tr(\cdot) + \left[(I-W)'X_1 + W'\overline{X}_2\right]'\left[\hat{V}\left(\hat{\beta}_1\right) + \hat{V}\left(\hat{\beta}_2\right)\right]\left[(I-W)'\overline{X}_1 + W'\overline{X}_2\right] \\ &+ \left(\hat{\beta}_1 - \hat{\beta}_2\right)'\left[(I-W)'\hat{V}\left(\overline{X}_1\right)(I-W) + W'\hat{V}\left(\overline{X}_2\right)W\right]\left(\hat{\beta}_1 - \hat{\beta}_2\right)\end{aligned} \qquad [3.23]$$

where Q is the explained part and U is the unexplained part.

CHAPTER 4

RESULTS AND IMPLICATIONS

Logit Model

If the assumptions of the model hold, the maximum likelihood (ML) estimators (e.g., the estimator produced by logit) are distributed asymptotically normally:

$$\hat{\beta}_k \overset{a}{\sim} N\left(\beta_k, \sigma^2{}_{\hat{\beta}_k}\right) \qquad [4.1]$$

The hypothesis $H_0 : \beta_k = \beta^*$ can be tested with the z-statistic:

$$z = \frac{\hat{\beta}_k - \beta^*}{\hat{\sigma}_{\hat{\beta}_k}}, \qquad [4.2]$$

where z is included in the computer output from the logit model. Under the assumptions justifying ML, if H_0 is true, then z is distributed approximately normally with a mean of zero and a variance of one for large samples.

Table 4.1 lists the logit model outcome for household variables, indicators related to mother, medical care variables, state Medicaid program variables, and number of times when influenza activity is widespread in 2000, 2001 and 2002 respectively (Wideflu00, Wideflu01 and Wideflu02). The state effects for the model are in table 4.2.

The model containing the same variables except for a modification in the influenza variable: number of times when influenza activity is widespread from 2000 to 2002(Wideflu3yr) is listed in table A.1 in the appendix. The state effects for this model are listed in table A.2.

Table 4.1. Logit Model for UTD Vaccination Status with Explanatory Variables [a]

Categories	UTD	Coefficient [b]	Standard Error	Z Value
Household variables	Agegrp	0.208*	0.021	10.150
	Wic	-0.076	0.049	-1.560
	Childnum	-0.222*	0.038	-5.850
	Frstbrn	0.072	0.047	1.530
	Incporat	0.078*	0.026	3.010
	Raceethk	-0.003	0.040	-0.060
	Sex	0.022	0.033	0.670
Indicators related to the mother	M_agegrp	0.096*	0.035	2.700
	Mumeduc	0.147*	0.041	3.570
	Marital	-0.007	0.048	-0.140
Medical care variables	Provnum	-0.196*	0.028	-7.120
	Compcare	-0.201*	0.027	-7.370
State Medicaid program variables	Medicaid	1.180*	0.223	3.240
	Separate	0.374**	0.168	2.220
Influenza level variables	Wideflu02	-0.022	0.021	-1.060
	Wideflu01	-0.012	0.018	-0.650
	Wideflu00	0.147*	0.021	7.120
	Intercept	-0.057	0.228	-0.250
Number of observations:	19,565			
Pseudo R^2:	0.037			
Log likelihood:	-11099.673			

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3.

[b] Significance levels *, **, and *** represent 1%, 5% and 10% significance level.

Table 4.2. Logit Model for UTD Vaccination Status with State Effects Only [a]

State	Coefficient [b]	Standard Error	Z Value
Alabama	0.882*	0.181	4.880
Alaska	-0.109	0.180	-0.600
Arizona	-0.097	0.167	-0.580
Arkansas	-0.238	0.206	-1.150
California	1.172*	0.189	6.200
Connecticut	1.930*	0.263	7.330
Delaware	0.400**	0.176	2.270
Dist. of Columbia	-0.032	0.199	-0.160
Florida	1.237*	0.193	6.410
Georgia	0.494*	0.149	3.320
Hawaii	0.158	0.198	0.790
Idaho	-0.927*	0.194	-4.770
Illinois	-0.539*	0.174	-3.100
Indiana	0.246	0.169	1.460
Kansas	0.422**	0.182	2.320
Kentucky	1.751*	0.244	7.190
Louisiana	-0.656*	0.163	-4.020
Maine	0.868*	0.221	3.930
Maryland	0.200	0.166	1.210
Massachusetts	1.072*	0.184	5.820
Michigan	0.955*	0.182	5.240
Minnesota	-0.544*	0.199	-2.730
Mississippi	1.321*	0.233	5.680
Missouri	-0.576*	0.221	-2.610
Montana	0.165	0.173	0.950
Nebraska	-0.490**	0.213	-2.300
Nevada	0.542*	0.187	2.900

Table 4.2. (Continued)

State	Coefficient	Standard Error	Z Value
New Hampshire	0.911*	0.217	4.210
New Jersey	0.434**	0.181	2.400
New York	0.281	0.226	1.240
North Carolina	0.513*	0.201	2.560
North Dakota	0.657*	0.216	3.040
Ohio	-0.846*	0.179	-4.730
Oklahoma	-0.019	0.188	-0.100
Oregon	0.806*	0.199	4.040
South Carolina	0.343	0.211	1.630
South Dakota	-0.772*	0.205	-3.760
Tennessee	-0.571*	0.199	-2.880
Texas	0.461**	0.209	2.210
Utah	0.347	0.193	1.800
Vermont	-0.250	0.156	-1.610
Virginia	0.510*	0.197	2.580
Washington	-0.204	0.122	-1.680
West Virginia	0.802*	0.228	3.520
Wisconsin	-0.556*	0.185	-3.000

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3.

[b] Significance levels *, **, and *** represent 1%, 5% and 10% significance level.

From the table 4.2, it is concluded that childhood vaccine compliance is different in different states, which is consistent with a previous research study by Stokley *et al*. Thirty-one states have significant coefficients at the 5% level: Alabama, California, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Kansas, Kentucky, Louisiana, Maine, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Nebraska, Nevada, New Hampshire, New Jersey, North Carolina, North Dakota, Ohio, Oregon, South Dakota, Tennessee, Texas, Virginia, West Virginia and Wisconsin. From the sign of the coefficients, further conclusion can be made that households' demand for childhood immunization will increase if they are located in Alabama, California, Connecticut, Delaware, Florida, Georgia, Kansas, Kentucky, Maine, Massachusetts, Michigan, Mississippi, Nevada, New Hampshire, New Jersey, North Carolina, North Dakota, Oregon, Texas, Virginia or West Virginia. On the other hand, households' demand for childhood vaccination will decrease if they are located in Idaho, Illinois, Louisiana, Minnesota, Missouri, Nebraska, Ohio, South Dakota, Tennessee or Wisconsin. According to the census regions and divisions of the U.S., they can be grouped into Northeastern, Midwestern, Southern, and Western states. There are five states: Connecticut, Maine, Massachusetts, New Hampshire, and New Jersey that belong to the Northeast. There are ten states: Illinois, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota and Wisconsin that belong to the Midwest. Twelve states: Alabama, Delaware, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Tennessee, Texas, Virginia and West Virginia that belong to the South. There are four states: California, Idaho, Nevada and Oregon belong to the West. From table 4.3, which lists the population and number of UTD households to population ratio for each state and each region, one can further study the regional difference among the 31 states. Texas holds the largest population and the most UTD households but its UTD percentage is not the

Table 4.3. Population and Up-to-Date Status for States

		Respondents		Up-to-date households		
Regions	**State**	Number	Percentage [a]	Number	Percentage [b]	State Percentage [c]
Northeast	Connecticut	256	1.31%	233	1.14%	87.02%
	Maine	256	1.31	179	0.91	69.47
	Massachusetts	519	2.65	433	2.21	83.40
	New Hampshire	248	1.27	195	1.00	78.74
	New Jersey	474	2.42	318	1.63	67.36
	Total	1753	8.96	1358	6.89	76.89
Midwest	Illinois	462	2.36	333	1.70	72.03
	Kansas	228	1.17	145	0.74	63.25
	Michigan	517	2.64	377	1.93	73.11
	Minnesota	241	1.23	180	0.92	74.80
	Missouri	267	1.36	199	1.02	75.00
	Nebraska	255	1.30	181	0.93	71.54
	North Dakota	257	1.31	163	0.83	63.36
	Ohio	710	3.63	519	2.65	73.00
	South Dakota	238	1.22	156	0.80	65.57
	Wisconsin	472	2.41	350	1.79	74.27
	Total	3647	18.63	2603	13.31	71.44
South	Alabama	490	2.50	393	2.01	80.40
	Delaware	266	1.36	196	1.00	73.53
	Florida	762	3.89	574	2.93	75.32
	Georgia	506	2.59	380	1.94	74.90
	Kentucky	252	1.29	211	1.08	83.72
	Louisiana	516	2.64	343	1.75	66.29
	Mississippi	232	1.19	185	0.95	79.83
	North Carolina	251	1.28	196	1.00	78.13
	Tennessee	828	4.23	627	3.20	75.65
	Texas	1307	6.68	900	4.60	68.86
	Virginia	267	1.36	209	1.07	78.68
	West Virginia	210	1.07	142	0.73	68.22
	Total	5887	30.08	4356	22.26	74.00

Table 4.3. (Continued)

		Respondents		Up-to-date households		
Regions	**State**	Number	Percentage	Number	Percentage	State Percentage
West	California	975	4.98	724	3.70	74.30
	Idaho	237	1.21	144	0.74	61.16
	Nevada	254	1.30	169	0.86	66.15
	Oregon	246	1.26	181	0.93	73.81
	Total	1712	8.75	1218	6.23	71.20

[a] Percentage represents the number of state respondents divided by the total number of respondents to the survey.

[b] Percentage represents the number of households in a state with up-to-date vaccine status divided by the total number of respondents to the survey.

[c] Percentage represents the number of households in a state with up-to-date vaccine status divided by the number of respondents within the state.

largest among the states. Connecticut has the highest UTD percentage among these 31 states. One can also see that the Northeast has the second smallest population among the four regions but has the highest UTD percentage. In contrast, western states have the smallest population but also have the lowest UTD percentage. States in the Northeast, South, and west regions with significant coefficients generally have a positive impact on vaccination up-to-date. Exceptions are Louisiana, Idaho and Tennessee. In contrast, states in the Midwest region generally have a negative impact on vaccinations. Exceptions are Kansas, Michigan and North Dakota. Possible explanation here is the different government policies.

From table 4.1, it can be concluded among the household variables, age of child (Agegrp), number of children in household (Childnm), and income to poverty ratio (Incporat) have a significant relationship with UTD status for childhood vaccination at the 1% significance level. From the sign (positive or negative) of the coefficients, it can be further concluded that households' demand for children's vaccine will increase as children grow from 19 to 35 months old. Households' compliance in childhood vaccination decreases as the number of children increases in the households. Households with higher income tend to keep their children's vaccine status more updated than their counterparts. Recalling the cross-tab provided in table 3.3, the conclusion regarding the relationship between the UTD status and age of children is confirmed by the outcome of the logit model. The logit result regarding the link between the UTD status and number of children in the household is also consistent with the cross-tab provided in table 3.3. This research outcome regarding the relationship between the age of child and the vaccination rate is consistent with Browngoehl, *et al*. They determined that as children grow from 20 to 35 months, their parents tend to give their children additional vaccines. Parents are less concerned about any serious side effects or reactions that their children may receive. The conclusion drawn about the number of children and the income of the households is consistent with the hypothesis that related to other households, households with higher income and fewer dependents can keep their vaccine status more up-to-date.

The relationship between the UTD status and the indicators related to the mother are also listed in table 4.1. From the table, age of mother (M_agegrp) and education level of mother (Mumeduc) have a significant relationship with UTD status for childhood vaccination at the 1% significance level. From the sign of the coefficients, it can further be concluded that households' demand for children's vaccine will increase as mothers' age increases from less than 19 to over 30 years old. Households with higher educated mother have higher vaccine rates than their lower educated counterparts. This conclusion is consistent with Kenkel (1990)'s who studied the relationship between consumer information and the demand for health care. Kenkel found that more informed individuals are more likely to keep their vaccine status up-to-date than poorly informed ones. Informed individuals realize the risk undertaken if they choose not to receive a vaccine. The reason why vaccine demand increases as mothers' age increases is complicated. One unlikely explanation is that women's immune system become less efficient as they become older. A more plausible explanation is with age comes wisdom.

Next consider the relationship between the UTD status and the medical care variables. From table 4.1, both variables in this category, number of providers (Provnum) and providers providing comprehensive childcare (Compcare) have a significant relationship with UTD status for childhood vaccination at the 1% significance level. From the sign of the coefficients, it can be concluded that as the number of providers increase, households' demand for childhood immunization decreases. Also, as more providers can offer comprehensive childcare in the household, it increases households' up-to-date vaccine status. This category is consistent with the cross tabs in table 3.3. It is also consistent with the research outcome by Joyce (2003) that a single private provider or one that provides comprehensive pediatric services makes household vaccination coverage higher than others.

Table 4.1 also lists the relationship between the UTD status and the state Medicaid program variables. Both variables in this category, Medicaid program (Medicaid) and separate new insurance program (Separate) have a significant relationship with UTD status for childhood vaccination at the 5% significance level. From the signs of the coefficients, it is concluded that participation in a Medicaid program or a new insurance program will increase households' demand for vaccine. This conclusion is consistent with Losasso's research outcome toward the effect of the insurance program.

Finally consider the relationship between the UTD status and the influenza level variables. Inspired by Mullahy (1998), who used a similar variable, individual's perceive risk of being infected by flu, it was expected the results would support the occurrence of vaccine-preventable diseases can stimulate vaccine demand. From table 4.1, only number of times when influenza activity is widespread in 2000 (Wideflu00) is a significant factor on deciding households' decision on childhood immunization at the 1% significance level. This is consistent with Philipson (1995). From the sign of coefficient, it is concluded that the more influenza takes place, the increased likelihood that people keep their vaccine status up-to-date. Unfortunately, from table 4.1, no significant relationship between the UTD status and the influenza variable Wideflu02 and Wideflu01 is found. The significance of Wideflu00 in the model is most likely spurious correlation.

After obtaining the results from logit models, the goodness of fit is evaluated. Summary statistics are displayed in table 4.4. The table indicates both models predict well. In the survey 27.6% of households are not up-to-date versus the model results predicting 26.6% are not. The confidence interval for those households is also very narrow: 26.0%-27.3%. On the other hand, 72.4% of households are not up-to-date versus the model results predicting 73.4% are not. The confidence interval for them is also narrow: 72.7%-74%.

Veall and Zimmermann (1992) provided interpretations of prediction-realization table and weaknesses. They found that using the percentage of correct predictions as their goodness-of-fit measure can give a misleading impression as to the explanatory value of the empirical model if one of the outcomes is particularly likely. They presented a “normalized” measure, which can be expressed by equation [4.3]:

$$\sigma_n = \sigma / \left(1 - p_{.1}^{\ 2} - p_{.2}^{\ 2}\right) = 4.810 \qquad [4.3]$$

where $p_{.i}$ is the fraction of times alternative I is predicted. In this case, $p_{.1} = 26.6\%$, $p_{.2} = 73.4\%$.

Veall and Zimmermann (1992) restricts attention to goodness-of-fit measures based on the prediction-realization table. Such measures have the weakness that strength of predictions is not taken into account.

Marginal Effect

In economics, the marginal effect or change is commonly used. The marginal effect is the change in predicted probability associated with changes in the explanatory variables. It is the partial derivative of the predicted probability/rate with respect to a given independent variable. The marginal effects are nonlinear functions of the parameter estimates and the levels of the

Table 4.4. Summary Statistics for Model Predictions

UTD	**Data**	**Predictions**	**Confidence Interval**
0	27.6%	26.6%	26.0 – 27.3%
1	72.4	73.4	72.7 – 74.0

[a] The level of confidence is 95%.

explanatory variables, so they cannot generally be inferred directly from the parameter estimates. The formula for expressing the marginal effect is:

$$\text{Marginal Change} = \frac{\partial \Pr(y = 1|x)}{\partial x_k} \qquad [4.4]$$

The marginal change is shown by the tangent to the probability curve in Figure 4.1. Since the dependent variable is a dummy variable, it is assumed that the relationship between those explanatory variables and the up-to-date vaccine status is essentially nonlinear. Given the nonlinearity of the model, the discrete change is used in the predicted probabilities for a given change in an independent variable. To define discrete change, two quantities are required:

$\Pr(y = 1|x, x_k)$ is the probability of an event given x, noting in particular the value of x_k.

$\Pr(y = 1|x, x_k + \delta)$ is the probability of the event with only x_k increased by some quantity δ.

Then, the discrete change for a change of δ in x_k equals

$$\frac{\partial \Pr(y = 1|x)}{\partial x_k} = \Pr(y = 1|x, x_k + \delta) - \Pr(y = 1|x, x_k) \qquad [4.5]$$

which can be interpreted as: a change in the predicted probability of an event changes by $\frac{\Delta \Pr(y = 1|x)}{\Delta x_k}$, given a change in x_k from x_k to $x_k + \delta$, holding all other variables constant.

As shown in Figure 4.1, $\frac{\partial \Pr(y = 1|x)}{\partial x_k} \neq \frac{\Delta \Pr(y = 1|x)}{\Delta x_k}$. They differ because the marginal change is the instantaneous rate of change, while the discrete change is the amount of change in the probability for a given finite change in one independent variable. The value of the discrete change depends on the starting level of the variable that is being changed, the amount of change

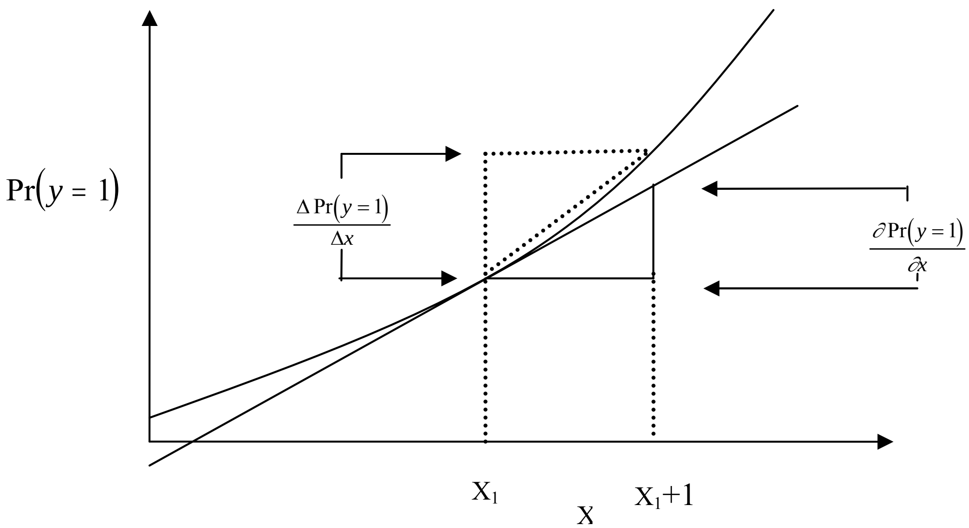

Figure 4.1. Marginal Change and Discrete Change in the Binary Regression Model

in that variable, and the level of all other variables in the model. For our model, discrete change is used with all variables held at their mean.

Table 4.5 displays the marginal effect for variables which are statistically significant at 5% level. First consider the marginal effect for the household variables. A 4% increase in UTD in childhood immunization will take place if a child grows from one lower age category to the adjacent higher age category. For each additional child in the household, the probability the household is up-to-date in vaccine status will decrease by 4.3%. Also, one unit increase in income to poverty ratio will increase the predicted probability by 1.5%.

Regarding the indicators related to mother, increasing age of mother from one lower category to the next higher age category increases 2% of the predicted probability. If a mother has some college, she will increase the UTD possibility of her household by 3% compared to her non college graduate counterpart.

The marginal effect for the medical care variables are also listed in table 4.5. Any additional provider will decrease the predicted probability of being up-to-date in vaccination by 4%. Those households that have multiple providers are lower in UTD vaccine coverage by 4% compared to those that only has one provider. This outcome is consistent with Joyce (2003).

Regarding the state Medicaid program variables, attending the Medicaid program will increase the UTD coverage in a large extent (20.8%) while attending the separate insurance program can also increase the UTD coverage by 7%.

Finally for the marginal effect for the influenza level variable, a one unit increment of widespread influenza during year 2000 will increase the probability of being up-to-date in childhood vaccination by 3%. Although the outcome is consistent with research by Philipson (1995)'s finding regarding the prevalence elasticity. The results indicate a very weak at

Table 4.5. Probability Change of Significant Independent Variables [a]

Categories	Variable	Marginal Effect [b]	Standard Error
Household variables	Agegrp	0.040*	0.004
	Childnum	-0.043*	0.007
	Incporat	0.015*	0.005
Indicators related to the mother	M_agegrp	0.019*	0.007
	Mumeduc	0.029*	0.008
Medical care variables	Provnum	-0.038*	0.005
	Compcare	-0.039*	0.005
State Medicaid program variables	Medicaid	0.208*	0.033
	Separate	0.070**	0.030
Influenza level variable	Wideflu00	0.029*	0.004

[a] Variables are statistically significant at 5% level in the model.

[b] Marginal effect=dy/dx. For the dummy variable, the marginal effect is the discrete change. Significance levels *, **, and *** represent 1%, 5% and 10% significance level.

best relation. Similar to Philipson's data selection, total number of measles cases reported by CDC's Morbidity and Mortality Weekly Report for the years 1984-91, the total number of widespread influenza activities reported from the Weekly Surveillance Reports of Influenza Branch at CDC from 2000-2002 is used. Philipson combined the prevalence levels by state with individual-level data on vaccine demand from the Child Health Supplement of the 1991 National Health Interview Survey (NHIS), which gives us the inspiration to use the state effect variables to better study the link between the influenza prevalence and the vaccine demand. Prevalence elasticity measures how level of infection affects the demand for vaccination. Philipson proved the significance and robustness of the estimated prevalence elasticities. He concluded that the time-varying covariate has a large and significant effect on duration to vaccination. Regardless of the specification used, the estimated effects are positive and highly significant at any standard level of significance. The consistency between our research outcome and Philipson's conclusion indicates our results regarding the influenza level variable are valid. Influenza outbreak does have positive effect on UTD status. Mullahy (1998) also had the similar findings: While it is costly (in terms of time costs) for workers to obtain immunization, workers who are more likely to get sick with the flu are more likely to have vaccination. Mullahy conclude that the extent to which individuals' perceived risks of infection may affect their propensities to be immunized. Mullahy's outcome is limited because his finding is only for flu shots by the elderly according to the significance of his regression results.

Odds Ratio

Effects for the logit model can be interpreted in terms of odds ratio. Recall that log of the odds is called the logit and that the logit model is linear in the logit, meaning that the log odds are a linear combination of the x's and β's. For example, consider a logit model with three independent variables:

$$\ln\left[\frac{\Pr(y=1|x)}{1-\Pr(y=1|x)}\right] = \ln\Omega(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \qquad [4.6]$$

For a unit change in x_k, logit changes by β_k, holding all other variables constant. The interpretation does not depend on the level of the other variables in the model. The problem is that a change of β_k in the log odds has little substantive meaning. Alternatively, by taking the exponential of both sides of this equation, creates a model that is multiplicative instead of linear, and results in an outcome which is a more intuitive measure, the odds:

$$\Omega(x, x_2) = e^{\beta_0} e^{\beta_1 x_1} e^{\beta_2 x_2} e^{\beta_3 x_3} \qquad [4.7]$$

where we consider the value of x_2. If letting x_2 change by 1,

$$\Omega(x, x_2+1) = e^{\beta_0} e^{\beta_1 x_1} e^{\beta_2 (x_2+1)} e^{\beta_3 x_3} = e^{\beta_0} e^{\beta_0} e^{\beta_1 x_1} e^{\beta_2 x_2} e^{\beta_2} e^{\beta_3 x_3} \qquad [4.8]$$

which leads to the odds ratio:

$$\frac{\Omega(x, x_2+1)}{\Omega(x, x_2)} = \frac{e^{\beta_0} e^{\beta_1 x_1} e^{\beta_2 x_2} e^{\beta_2} e^{\beta_3 x_3}}{e^{\beta_0} e^{\beta_1 x_1} e^{\beta_2 x_2} e^{\beta_3 x_3}} = e^{\beta_2} \qquad [4.9]$$

Accordingly, the exponential of the coefficient is interpreted as: for a unit change in x_k, the odds are expected to change by a factor of $\exp(\beta_k)$, holding all other variables constant.

For $\exp(\beta_k) > 1$, the odds are "$\exp(\beta_k)$ times larger". For $\exp(\beta_k) < 1$, the odds are "$\exp(\beta_k)$ times smaller". The effect of a standard deviation change in x_k instead of a unit change is: for a standard deviation change in x_k, the odds are expected to change by a factor of $\exp(\beta_k \times s_k)$, holding all other variables constant.

The outcome of the odds ratio for both a unit and a standard deviation change of the independent variables is provided in table 4.6. Since the results for unit change in odds ratio and the results for standard deviation change in odds ratio are very similar, only interpretation for the unit change in odds ratio is presented here. Regarding the change in odds ratio for household variables, the outcome can be interpreted as: For every increase in age category of child, the odds of being UTD in vaccine status increase by a factor of 1.231-1=0.231, holding other variables constant. For each additional child in the household, the odds of being UTD in vaccine status decrease by a factor of .2, holding other variables constant. One unit increase in income to poverty ratio will increase the odds of being UTD in vaccine status by a factor of 0.081.

Similarly, table 4.6 lists the change in odds ratio for indicators to the mother. For every increase in age category of mother, the odds of being UTD in vaccine status increase by a factor of 0.1, holding other variables constant. Mother having some college will increase the odds of being UTD in vaccine status by a factor of .16, holding other variables constant.

Regarding the medical care variables, the table indicates for a unit increase in the number of providers, the odds of being UTD in vaccine status decrease by a factor of 0.18, holding all other variables constant. For a unit increase in the number of providers providing comprehensive childcare in household, the odds of being UTD in vaccine status increase by a factor of 0.19, holding all other variables constant.

Similarly, the table lists the change in odds ratio for state Medicaid program variables. Participation in the Medicaid program increases the odds of being UTD in vaccine status increase by a factor of 2.26, holding other variables constant. On the other hand, participating in the separate insurance only increases the odds by a factor of 0.45.

Table 4.6. Change in Odds Ratio for Significant Independent Variables [a]

Categories	Variable Name	Standard Deviation	Factors Change in Odds	
			Unit	Standard Deviation
Household variables	Agegrp	0.803	1.231	1.182
	Childnum	0.604	0.801	0.874
	Incporat	0.961	1.081	1.078
Indicators related to the mother	M_agegrp	0.537	1.100	1.053
	Mumeduc	0.494	1.158	1.075
Medical care variables	Provnum	0.594	0.822	0.890
	Compcare	0.583	0.818	0.890
State Medicaid program variables	Medicaid	0.474	3.257	1.750
	Separate	0.435	1.453	1.176
Influenza level variable	Wideflu00	2.311	1.158	1.403

[a] Variables are statistically significant at the 5% level in the model.

Finally, a unit increase in the number of widespread influenza taking place will increase the odds of being UTD in vaccine status increase by a factor of 0.15, holding other variables constant.

Oaxaca-Blinder Decomposition

For the Oaxaca-Blinder decomposition, table 4.7 lists regression results for all the variables in the model including the coefficients for household variables, indicators related to mother, medical care variables, state Medicaid program variables and number of times when influenza activity is widespread in 2000, 2001 and 2002 for both the lower-educated mother and the higher-educated mother. Higher educated mother group here refers to those who have had some college while lower educated mother group refers to those who have not had any college. Regression results for the model that contains different influenza information are provided in the appendix in table A.3. State effect variables are included in the model. The state effect coefficients for the model are provided in Appendix tables A.4 and the state effect coefficients for the model that contains different influenza information are listed in Appendix table A.5.

Table 4.7 displays coefficients for both the high-educated mother group and low-educated mother group for the two logit models. Significant differences between the intercepts of the state effects of the higher-educated and lower-educated mother groups are tested. Thirty-six out of 49 the states in the t-test indicate no significant difference between the intercepts at the 5% significance level. In general, it will be assumed for the Oaxaca-Blinder decomposition (later in this chapter) that the overall intercepts between higher-educated and lower-educated mother groups are not significantly different.

Based on table 4.7, there are two groups of households which have different mother's education level (college or no college). The Oaxaca-Blinder decomposition indicates: How much

Table 4.7. Regression Results for Higher- and Lower-Educated Mother Groups [a]

Variable Names	Coefficients for Mothers [b]	
	Higher-educated mothers	Lower-educated mothers
Agegrp	0.216*	0.205*
Wic	-0.142	-0.035
Childnum	-0.308*	-0.184*
Frstbrn	-0.111	0.186
Incporat	0.149*	0.050*
Marital	0.031	0.012
M_agegrp	0.137**	0.079**
Raceethk	-0.064	0.042
Sex	-0.017	0.044
Provnum	-0.237*	-0.176*
Compcare	-0.175*	-0.208*
Medicaid	-0.251	-0.807
Separate	-0.683	-0.882
Wideflu02	-0.017	-0.095
Wideflu01	-0.043	0.081
Wideflu00	0.110*	-0.041*
Intercept	1.512	1.791
Number of observations	8266	11299
Pseudo R^2	0.039	0.034
Log likelihood	-4322.200	-6727.546

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3

[b] Significance levels *, **, and *** represent 1%, 5% and 10% significance level.

of the gap between vaccines status of these two educational levels is due to difference in household endowments between mothers with and without college and difference in the estimated coefficients (marginal impact of explanatory variables). The decomposition decomposes the effect that mother's education has on UTD into the direct and indirect effects. The direct effect is where the education level directly changes the level of UTD status. The indirect effect is where the education level changes the level of other explanatory variables (say income) and thus changes the level of UTD status through the explanatory coefficients.

Table 4.8 lists the results of decomposition for households with mother having some college versus the households with mothers without any college. The estimates for the endowment effects are calculated as $\beta_i^H \cdot \left(\overline{X}_i^H - \overline{X}_i^L\right)$, where β_i^H is the coefficient associated with explanatory variables, i for high-educated mothers in table 4.7. $\overline{X}_i^H$ and $\overline{X}_i^L$ are mean values of the explanatory variable. Similarly, the estimates for the coefficient effects are calculated as $\overline{X}_i^L \cdot \left(\beta_i^H - \beta_i^L\right)$, where β_i^L is the coefficient associated with explanatory variable (i for low-educated mother in table 4.7). Summing the endowment estimates over the entire explanatory variables (including the state effects) yield the direct effect called endowment effect,

$$E = \sum_{i=1}^{n} \left(\beta_i^H \cdot \left(\overline{X}_i^H - \overline{X}_i^L\right)\right) = 36.0\% \qquad [4.10]$$

where n is the number of explanatory variables including the state effects. Summing the coefficient estimates over all the explanatory variables yields the indirect effect called the coefficient effect,

$$C = \sum_{i=1}^{n} \left(\overline{X}_i^L \cdot \left(\beta_i^H - \beta_i^L\right)\right) = 19.3\% \qquad [4.11]$$

The total effect called the amount attributable of education on UTD is $E + C = 55.3\%$. This

Table 4.8. Decomposition Results for Independent Variables [a]

	Effects			
	Characteristics [b]		Coefficients Effects [c]	
Categories	**Estimate**	**Share**	**Estimate**	**Share**
Household variables				
Agegrp	0.1	0.2%	2.2	4.0%
Wic	7.2	13.0	-7.3	-13.2
Childnum	3.9	7.1	-23.6	-42.7
Frstbrn	-0.8	-1.4	-10.9	-19.7
Incporat	15.1	27.3	16.2	29.3
Raceethk	1.7	3.1	-5.6	-10.1
Sex	0.0	0.0	-3.0	-5.4
Indicators related to the mother				
M_agegrp	5.5	9.9	13.7	24.8
Marital	0.7	1.3	1.4	2.5
Medical care variables				
Provnum	1.9	3.4	-8.7	-15.7
Compcare	1.5	2.7	4.2	7.6
State Medicaid program variables				
Medicaid	-0.9	-1.6	16.4	29.7
Separate	-0.1	-0.2	-2.5	-4.5
Influenza level variable				
Wideflu02	0.0	0.0	-3.4	-6.1
Wideflu01	-0.3	-0.5	-16.9	-30.6
Wideflu00	0.7	1.3	70.5	127.5
Total	36.0	65.1	19.3	34.9

[a] Statistically significant at the 5% level.

[b] Endowment effects explains the differences in the mean values of the explanatory variables

[c] Coefficients effects explains the difference in the estimated coefficients

percentage is the portion that education explains the difference in UTD for higher- and lower-educated mothers. The shares are determined by taking the estimates and dividing them by the amount attributable.

Approximately 65.1% of the amount attributable to education is explained by the endowment effect. This indicates that if a household with a non-college mother received some college, then the UTD gap would have been $(65.1\%)\cdot(55.3\%) = 36\%$ less. Differences in income to poverty ratio accounts for 27.3% of the amount attributable.

Thirty-four point nine percent of the amount attributable between households with an educated mother and non-educated are explained by the differences in logit coefficients. If in both groups the various variables influencing UTD status had the same strength (their coefficients in the logit equation had been equal), then about 34.9% of the increased probability of being UTD for households with educated mother would disappear.

The coefficient effect of the income to poverty ratio variable is large between educated and non-educated mother households, accounting for 29.3 percent of the difference in probability of the amount attributable. This suggests that for educated mother households, higher incomes have contributed to the greater incidence of UTD among such households due to the higher returns households received for the jobs it holds as compared to non-educated mother households.

Table 4.9 lists the summary of the decomposition results. The first three rows are the amount attributable, characteristic effect (E) and coefficient effect (C) in table 4.8.

As mentioned earlier in this chapter, t-tests indicate no significant difference between the intercepts at the 5% significance level. In general, it is assumed for the Oaxaca-Blinder decomposition that the overall intercepts between higher-educated and lower-educated mother groups are not significantly different and this is reflected in the 0.0% in U (the shift coefficient).

U = higher-educated mother intercept – lower-educated mother intercept [4.12]

Raw differential $R = E + C + U$ is the unadjusted difference in UTD among the households. The adjusted differential $D = C + U$ accounts for the coefficient effects(C) plus the shift coefficient. The endowment is a percent of total raw differential is E/R and preferences as percent of total is D/R.

The standard errors for the decomposition are not computed in the analysis because Jann's Oaxaca-Blinder decomposition program in STATA which calculates the standard errors is not available for the logit model. Alternative methods for computing the decomposition for binary limited dependent variable models have been proposed yet methods to compute standard errors for the decomposition have not received a high priority in applied research. Future extensions of the research will investigate the application of Yun's technique using STATA to develop standard errors. The standard errors will permit the evaluation of the statistical significance of the reported decomposition results.

Table 4.9. Summary of Decomposition Results [a]

Amount attributable:	55.3%
Due to characteristics (E):	36.0
Due to coefficients (C):	19.3
Shift coefficient (U):	0.0
Raw differential (R) {E+C+U}:	55.3
Adjusted differential (D) {C+U}:	19.3
Characteristics as % total (E/R):	65.1
Preferences as % total (D/R):	34.9

[a] E= portion due to characteristic effects

C= portion due to coefficient effects

U = unexplained portion of differential (difference between model constants)

D = portion due to preferences (C+U)

CHAPTER 5

SUMMARY AND CONCLUSION

Health workers in the U.S. warn that parents' refusal to have their children vaccinated could lead to widespread deadly diseases. It is hypothesized that because some deadly diseases have been wiped out in the U.S, some households have become complacent of their children's current vaccination status and become less aware of the importance to get their children up-to-date in vaccination. CDC has issued a report warning that if children are not vaccinated against serious diseases, then it could lead to a serious public health problem. CDC warns that vaccines for diseases such as measles and diphtheria can only be effective if given to a high proportion of children. CDC also states that the risk of side effects of vaccination is extremely low.

Results in table 4.1 indicate a strong correlation is found between some household variables and the households' demand for childhood vaccination. It can be concluded that childhood immunization compliance increases as the age of the children increases. As children age from 19 to 35 months, parents tend to give their children additional vaccines because they are less worried about any serious side effects that their children may experience from vaccination. It also indicates that households' compliance in childhood vaccination decreases as the number of children increases in households. Large number of children may cause financial hardship and time constraints for households and thus leads to the low vaccination coverage. Income of the households is a very traditional factor that affects households' childhood vaccine demand. High income households tend to keep their children more up-to-date in vaccination than their lower income counterparts.

Among the household variables, household income is an important factor to affect houshold's demand for childhood vaccines. One unit increase in the income to poverty ratio will lead to a 1.5% increase in UTD vaccine status. Therefore, cost effectiveness is crucial in government's policy toward childhood vaccination.

Indicators related to the mother of the household also have a large impact on households' compliance in childhood immunization. Considering the two mother groups with different educational levels, the higher educated mothers have a higher up-to-date vaccination rate (40%) than the lower educated (non-college) mothers (32%). Better educated individuals are found to use more vaccinations. It confirms the earlier findings of Hay and Leahy (1982). Information can increase the probability that the individual makes a physician visit. The number of physician visits has a positive correlation with households' demand for childhood immunization. It can be further concluded that up-to-date childhood immunization compliance increases with mothers' age.

The relationships among the households' demand for childhood immunization and the medical care variables are found to be statistically significant at the 1% significance level. Regression results indicate that half of the households that are up-to-date in their childhood vaccine status have only one provider. They also indicate that up-to-date childhood immunization rate is very high (63%) if all the providers of the child can provide comprehensive pediatric services.

State Medicaid program variables are found to be positive factors to affect households' compliance in childhood vaccination. For those who participated in the Medicaid program, over

73% of the households are found to be up-to-date in childhood vaccine status. For those who took part in the separate insurance program, 72% of the households are up-to-date. For those who took part in the combination program, 73% of the households are up-to-date. It also implies that there are more people participating the Medicaid program than purchasing separate insurance. As mentioned in chapter 4, the Medicaid program has a crowding-out effect on separate insurance program. Private insurance is negatively associated with Medicaid eligibility and has a positive impact on up-to-date vaccine status. There were large and statistically significant increases in Medicaid eligibility between 1988 and 1991. For example, the proportion of white children eligible for Medicaid rose from 20% to 36%. Among blacks, the proportion rose from 45% to 59% in the same period. On the other hand, private insurance coverage dropped for all three groups. For example, among whites, private insurance dropped from 77% to 71%. The declines in private insurance may reflect a decline in health benefits associated with low-wage employment during this period or displacement of privately purchased insurance or both. It is also noticed that over 40% of the households opted for the combination of the Medicaid program and separate insurance. One explanation is that households tend to minimize their cost and risk since insurance is negatively associated with Medicaid. Many other new programs appeared in these years to eliminate cost as a barrier and improve immunization levels such as State Children's Health Insurance Program (SCHIP) and Vaccine for Children (VFC).

Geographic factors may also play an important role in deciding households' demand for childhood vaccination. Results indicate households' demand for childhood immunization will increase if they are located in Alabama, California, Connecticut, Delaware, Florida, Georgia, Kansas, Kentucky, Maine, Massachusetts, Michigan, Mississippi, Nevada, New Hampshire, New Jersey, North Carolina, North Dakota, Oregon, Texas, Virginia or West Virginia. On the other hand, Households' demand for childhood vaccination will decrease if they are located in Idaho, Illinois, Louisiana, Minnesota, Missouri, Nebraska, Ohio, South Dakota, Tennessee or Wisconsin. Connecticut has the highest UTD percentage among these 31 states. If the U.S is divided into Northeast, Midwest, South and West, it can be further concluded that the Northeast has the second smallest population among the four regions but has the highest UTD percentage. In contrast, western states have the smallest population but also have the lowest UTD percentage.

A weak correlation is found between the extent of influenza activity and the households' demand for childhood vaccination. Results indicate a one unit increment of widespread influenza during year 2000 will increase the rate of being up-to-date in childhood vaccination by less than 3%.

Fully immunizing your child according to physician's recommendations can help protect child from many common infections. These infections can lead to serious and even life threatening complications. Although the vaccines may have mild side effects and more rarely, serious complications, in general, it is safer to immunize children than allow them to get any of the infections that they prevent.

Results indicate state Medicaid program improves the childhood vaccination coverage but only 70% of the households attend the Medicaid program. Therefore improvements in the Medicaid program for children are needed. A dearth of medical providers in low income urban and rural areas undermines the capability of Medicaid to realize improvements in utilization of health care services, and potentially health status, especially for minority children (Fosset *et al*., 1992). Furthermore, since most households have only one provider, particular attention should be paid to the issue of provider vaccine education for children newly eligible for coverage.

The major result from this research is the effect of household preferences toward up-to-date vaccine status. The Oaxaca-Blinder decomposition of mother's education indicates a marked difference between college and non-college mothers in household preferences toward up-to-date vaccine status. The percentage of difference in immunization compliance between college and non-college mothers due to household vaccine preference is 35%. The result supports the belief that a major cause of non-immunization compliance results from household preference. This difference in household preferences may result from customs, habits, fear, or lack of knowledge on the benefits of compliance. In any case, the results strongly support the hypothesis that changing these preferences of lower educated mothers can have a major positive impact on increasing the vaccination compliance rate. Targeting immunization education programs at these lower educated households, yielding a positive change in preferences, has a great potential of increasing the vaccination compliance rate. With this targeting the goal that 90% of children receive all recommended immunizations by age two can be achieved and maintained.

This study is subject to the number of limitations. There might be some misclassification of some recoded variables such as race or ethnicity, mother's education and WIC. The WIC variable is not statistically significant in the logit model, which conflicts with Morrow *et al.*'s previous research work. The lack of time series dates prevents investigation of the trends in vaccination compliance.

Risk/benefit analysis should be seriously considered in the future research work. Because when children receive immunizations at a health care facility, they also have a wide range of preventive services made available to them. Immunizations are only one measure of the overall delivery of all preventive health services to children.

Furthermore, the federal government should direct additional resources to addressing some of the remaining questions regarding publicly funded children's health insurance. Propaganda to the medical firms regarding the importance of vaccination is necessary because the firm can make arbitrarily higher revenue in percentage terms with drug treatments than with vaccines, according to Kremer and Snyder.

Obstacles to childhood vaccine availability — such as immunization expenses, weak markets, and difficulties in predicting need — often have economic roots. As mechanisms to enhance availability, government may consider financial incentives, improved coordination, and alternatives to safety and effectiveness documentation.

A pillar of U.S. policy on vaccines is the protection of the individuals who use them. FDA does not license a product for sale in the United States until it is satisfied that the vaccine is safe and effective. Scientists, clinicians, members of Congress, and the public must make decisions of vaccine safety despite uncertainties and varying perceptions of risk. To ameliorate the difficulties, government could address education and risk communication, studies in childhood vaccination and health economics, and improve available mechanisms to compensate individuals injured by vaccinations.

Regarding the household, government may also set a limitation on the number of providers that can provide comprehensive child care to the household. Noting concern for health needs of low-income household, development of affordable childhood vaccines for them is a good way to help spur the vaccination coverage.

REFERENCES

Browngoehl, K., K. Kennedy, K. Krotki, and H. Mainzer. "Increasing Immunization: A Medicaid Managed Care Model." *Pediatrics*. 99(1) 1997:4-8.

Davis, T.C., D.D. Fredrickson, C.L. Arnold, J.T. Cross, S.G. Humiston, K.W. Green, and J.A. Bocchini. "Childhood Vaccine Risk/Benefit Communication: A National Survey." *Pediatrics*. 107(2) 2001:17-27.

Dubay, L., and G. Kenney. "Expanding Public Health Insurance to Parents: Effects on Children's Coverage under Medicaid." *HSR: Health Services Research*. 38(5) 2003:1283-1302.

Dude, A.M. "Recent Trends in Infant Mortality In Kazakhstan: The Result of Discrimination?" International Union for the Scientific Study of Population XXV International Population Conference: University of Chicago, Feb 2005.

Feikema, S.M., R.M. Klevens, M.L. Washington, and L. Barker. "Extraimmunization Among US Children." *The Journal of the American Medical Association*. 283(10) 2000:1311-1317.

Fossett, J.W., J.D. Perloff, P.R. Kletke, and J.A. Peterson. "Medicaid and Access to Child Health Care in Chicago." *Journal of Health Politics, Policy and Law*. 17(2) 1992:273-298.

Gang, I.N., K. Sen, and M-S. Yun. "Caste, Ethnicity and Poverty in Rural India." Discussion Paper No. 629. Institute for the Study of Labor DP No. 629 <http://urban.hunter.cuny.edu/RePEc/seminar/Gang-Sen-Yun-HunterSeminar-30Sept05.pdf>

Hay, J.W., and M.I. Leahy. "Physician Induced Demand: An Empirical Analysis of the Consumer Information Gap." *Journal of Health Economics*. 1(3) 1982:231–244.

Jann, B. "Standard Errors for the Blinder-Oaxaca Decomposition." 3rd German Stata Users Group Meeting Berlin: Department of Sociology, ETH Zurich, April 2005. <http://www.stata.com/meeting/3german/jann.pdf>

Joyce, T., and A. Racine. "Chip Shots: Association Between The State Children's Health Insurance Programs And Immunization Coverage And Delivery." *Pediatrics*. 115(5) 2005:526-534.

Kenkel, D.S. “Consumer Health Information and the Demand for Medical Care.” *The Review of Economics and Statistics*. 72(4) 1990:587-595.

__________. “The Demand for Preventive Medical Care.” *Applied Economics.* 26(4) 1994:313-325.

Kenyon, T.A., M.A. Matuck, and G. Stroh. “Persistent Low Immunization Coverage Among Inner-city Preschool Children Despite Access to Free Vaccine.” *Pediatrics*. 101(4) 1998:612-616.

Kitagawa, E.M. “Components of A Difference Between Two Rates.” *Journal of the American Statistical Association*. 50(272) 1995:1168-1194.

Kremer, M., and C.M. Snyder. “Why Are Drugs More Profitable Than Vaccines?” NBER Working Paper Series No. w9833. Cambridge, MA: National Bureau of Economic Research.

Kutty, N.K. “The Demand for Home Modifications: A Household Production Function Approach.” *Applied Economics*. 31(10) 1999:1273–1281.

__________. “Production of Functionality by the Household: A Household Production Function Approach.” *Applied Economics*. 32(10) 2000:1269–1280. <http://adfdell.pstc.brown.edu/classes/readings/kut98.pdf>

Liao, T.F. *Interpreting Probability Models Logit, Probit, and Other Generalized Linear Model,* University of Illinois, Urbana-Champaign, 1994.

________. *Statistical Group Comparison.* Canada: John Wiley & Sons, Inc., 2002.

LoSasso, A.T., and T.C. Buchmueller. “The Effect of the State Children’s Health Insurance Program.” *Journal of Health Economics*. 23(5) 2004:1059–1082.

Madden, D. “Towards a Broader Explanation of Male-Female Wage Differences.” *Applied Economics Letters*. 7(12) 2000:765-770.

Manning, A. “Immunization Rate Exceeds Federal Goals.” *USA Today*, July 27, 2005.

McGuire, T. “Setting Prices for New Vaccines (in Advance).” *International Journal of Health Care*. 3(3) 2003:207-224.

Morrow, A.L., J. Rosenthal, H.D. Lakkis, J.C. Bowers, F.D. Butterfoss, R.C. Crews, and B. Sirotkin. “A Population-based Study of Access to Immunization Among Urban Virginia Children Served By Public, Private, and Military Health Care Systems.” *Pediatrics*. 101(2) 1998:5-16.

Mullahy, J. “It’ll Only Hurt A Second Microeconomic Determinants of Who Gets Flu Shots.” *Health Economics*. 8(1) 1999:9-24.

Oaxaca, R., and M.R. Ransom. “On Discrimination and the Decomposition of Wage Differentials” *Journal of Econometrics*. 61 1994:5–21.

Pauly, M.V., and B.E. Cleff. “The Economics of Vaccine Policy.” *In Supplying Vaccines: An Economic Analysis of Critical Issues.* Pauly, M.V., C.A. Robinson, S.J. Sepe, M. Sing, M.K. Willian. (eds). Burke, VA: IOS Press, Inc, 1996.

Philipson, T. “Private Vaccination and Public Health: An Empirical Examination for U.S. Measles.” *The Journal of Human Resources*. 31(3) 1995:611-630.

Richards, S., and L. Stemnock. “Estimated Influenza Vaccine Coverage Indiana and United States, January 2005.” *India Epidemiology Newsletter*, February 2005, VOL.XIII, No. 2.

Silberman, J., D.A. Gerlowski, and N.A. Williams. “Estimating Existence Value for Users and Nonusers of New Jersey Beaches.” *Land Economics.* 68(2) 1992:225–236.

Stokley, S., P.J. Smith, R.M. Klevens, and M.P. Battaglia. “Vaccination Status of Children Living in Rural Areas in the United States. Are They Protected?” *American Journal of Preventive Medicine*. 20(4S) 2001:55-60.

Veall, M.R., and K.F. Zimmermann. “Performance Measures from Prediction-Realization Tables.” *Economics Letters* 39 1992:129-134

Winsborough, H.H., and P. Dickinson. *Components of Negro-White Income Differences* Proceedings of the Social Statistics, Section:6-8, 1971.

Yun, M-S. “Decomposing Differences in the First Moment.” *Economics Letters*. 82(2) 2004: 275-280.

APPENDICES

Table A.1. Logit Model for UTD Vaccination Status with Explanatory Variables [a]

Categories	UTD	Coefficient [b]	Standard Error	Z Value
Household variables	Agegrp	0.208*	0.021	10.150
	Wic	-0.076	0.049	-1.560
	Childnum	-0.222*	0.038	-5.850
	Frstbrn	0.072	0.047	1.530
	Incporat	0.078*	0.026	3.010
	Raceethk	-0.003	0.040	-0.060
	Sex	0.022	0.033	0.670
Indicators related to the mother	M_agegrp	0.096*	0.035	2.700
	Mumeduc	0.147*	0.041	3.570
	Marital	-0.007	0.048	-0.140
Medical care variables	Provnum	-0.196*	0.028	-7.120
	Compcare	-0.201*	0.027	-7.370
State Medicaid program variables	Medicaid	0.721*	0.223	3.240
	Separate	-0.429**	0.197	-2.180
Influenza level variables	Wideflu3yr	0.015	0.009	1.640
	Intercept	0.746*	0.212	3.510
Number of observations:	19565			
Pseudo R^2:	0.037			
Log likelihood:	-11099.673			

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3.

[b] Significance levels *, **, and *** represent 1%, 5% and 10% significance level.

Table A.2. Logit Model for UTD Vaccination Status with State Effects Only [a]

State	Coefficient [b]	Standard Error	Z Value
Alabama	0.604*	0.192	3.150
Alaska	-0.709*	0.226	-3.140
Arizona	0.301**	0.147	2.050
Arkansas	-0.187	0.227	-0.830
California	0.369**	0.169	2.190
Connecticut	1.046*	0.242	4.330
Delaware	0.663*	0.183	3.620
Dist. of Columbia	-0.375	0.223	-1.680
Florida	0.434**	0.173	2.510
Georgia	0.757*	0.157	4.820
Hawaii	-0.185	0.223	-0.830
Idaho	-0.876*	0.216	-4.060
Illinois	-0.488**	0.198	-2.470
Indiana	-0.031	0.181	-0.170
Iowa	-0.146	0.209	-0.700
Kansas	0.310	0.179	1.740
Kentucky	0.949*	0.228	4.160
Louisiana	-0.737*	0.191	-3.850
Maine	0.065	0.203	0.320
Maryland	-0.208	0.200	-1.040
Massachusetts	0.794*	0.195	4.070
Michigan	0.416**	0.182	2.280
Minnesota	-0.619**	0.250	-2.480
Mississippi	0.700*	0.228	3.080
Missouri	-0.412	0.233	-1.760
Montana	0.296	0.178	1.670
Nebraska	-0.626*	0.229	-2.740

Table A.2. (Continued)

State	Coefficient	Standard Error	Z Value
Nevada	0.542*	0.187	2.900
New Hampshire	0.428	0.221	1.940
New Jersey	-0.029	0.186	-0.150
New Mexico	-0.492**	0.222	-2.220
New York	0.110	0.231	0.470
North Carolina	0.903*	0.180	5.020
North Dakota	-0.145	0.199	-0.730
Ohio	-0.533*	0.192	-2.780
Oklahoma	-0.443**	0.219	-2.030
Oregon	0.752*	0.187	4.010
Pennsylvania	0.976*	0.144	6.760
Rhode Island	-0.220	0.273	-0.810
South Dakota	-0.870*	0.225	-3.860
Tennessee	-0.587**	0.262	-2.240
Texas	0.054	0.197	0.270
Utah	0.532*	0.194	2.750
Vermont	0.126	0.161	0.780
Virginia	0.668*	0.201	3.320
Washington	0.061	0.131	0.460
Wisconsin	-0.485**	0.207	-2.350

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3.

[b] Significance levels *, **, and *** represent 1%, 5% and 10% significance level.

Table A.3. Regression Results for Higher-and Lower-Educated Mother Groups [a]

Variable Name	Coefficients for Mothers [b]	
	Higher-Educated mothers	Lower-Educated mothers
Agegrp	0.216*	0.205*
Wic	-0.142	-0.035
Childnum	-0.308*	-0.184*
Frstbrn	-0.111	0.186*
Incporat	0.149*	0.050
Marital	0.031	0.012
M_agegrp	0.137**	0.079
Raceethk	-0.064	0.042
Sex	-0.017	0.044
Provnum	-0.237*	-0.176*
Compcare	-0.175*	-0.208*
Medicaid	-1.165*	0.766*
Separate	-0.778**	0.104***
Wideflu3yr	-0.001	0.040*
Intercept	1.607	0.415
Number of observations	8266	11299
Pseudo R^2	0.039	0.034
Log likelihood	-4322.200	-6727.546

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3. Higher educated mother group here refers to those who have ever had some college. Lower educated mother group refers to those who have not had any college.

[b] Significance level *, **, and *** represent 1%, 5% and 10% significance level.

Table A.4. Logit Model for UTD Vaccination Status with State Effects Only for Different Mother Groups [a]

State	Coefficients for Mothers [b]	
	Higher-Educated mothers	Lower-Educated mothers
Alabama	-0.579	-0.251
Alaska	-0.447	0.057
Arizona	-0.202	0.754*
Arkansas	-0.122	0.319
California	-0.407	-0.733*
Delaware	0.277	-0.227
Dist. of Columbia	0.511	0.197
Florida	-0.306	-0.601
Georgia	-0.437	0.174
Hawaii	0.112	0.215
Idaho	0.077	-0.169
Illinois	-1.151*	0.044
Indiana	-0.476	-1.030*
Iowa	-0.958*	-1.221*
Kansas	-1.010*	-0.326
Kentucky	0.447	-0.332
Louisiana	0.580	-0.344
Maine	-0.478	-0.881**
Maryland	-0.911**	0.141
Massachusetts	0.276	-0.167
Michigan	-0.224	-0.683*
Minnesota	-0.332	0.141
Mississippi	-0.340	-0.528
Missouri	-0.294	0.596**
Montana	-0.557	-0.105

Table A.4. (Continued)

State	Coefficients for Mothers	
	Higher-Educated mothers	Lower-Educated mothers
Nebraska	-0.398	0.431
Nevada	-0.327	-0.114
New Hampshire	-0.874**	-0.438
New Jersey	-0.873	-1.149*
New Mexico	0.596	0.285
North Dakota	-0.722**	-1.376*
Ohio	-0.587	0.030
Oklahoma	-0.061	-0.221
Oregon	0.385	-0.051
Pennsylvania	0.433	0.895**
South Carolina	-0.578	0.291
Tennessee	-0.082	0.500
Texas	-0.804	-0.101
Utah	0.390	0.714*
Vermont	-0.775*	0.309
Virginia	0.256	0.995*
Washington	-0.377	-0.452
West Virginia	-0.937**	-0.978*
Wisconsin	-0.528	0.490**
Wyoming	-0.450	-0.655

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3.

[b] Significance level *, **, and *** represent 1%, 5% and 10% significance level.

Table A.5. Logit Model that contains different influenza information for UTD Vaccination Status with State Effects Only for Different Mother Groups [a]

State	Coefficients for Mothers	
	Higher-Educated mothers	Lower-Educated mothers
Alabama	-0.232	0.799*
Alaska	-0.594	-0.520
Arizona	0.229	-0.483**
Arkansas	0.297	-0.122
California	-0.503	0.643*
Colorado	0.056	-0.873*
Connecticut	0.732**	1.496*
Florida	-0.219	0.774*
Georgia	-0.533	0.400
Hawaii	0.333	0.018
Idaho	0.165	-0.611**
Illinois	-0.731	-0.398
Indiana	-0.057	0.020
Iowa	-0.611	-0.253
Kansas	-0.553	-0.342
Kentucky	0.398	1.043*
Louisiana	0.485	-0.704*
Maine	-0.169	0.495
Maryland	-1.007*	-0.151
Massachusetts	0.306	0.883*
Michigan	0.123	0.530**
Minnesota	-0.206	-0.763**
Mississippi	-0.044	0.804*
Missouri	0.281	-0.550
Montana	-0.446	0.203

Table A.5. (Continued)

State	Coefficients for Mothers	
	Higher-Educated mothers	Lower-Educated mothers
Nebraska	-0.061	-0.686**
Nevada	-0.124	0.275
New Hampshire	-0.721**	0.422
New Jersey	0.021	0.062
New Mexico	-0.352	-0.453
New York	0.829**	-0.102
North Carolina	-0.817**	0.264
Ohio	0.053	-0.575**
Oklahoma	-0.099	-0.298
Oregon	0.301	0.419
Pennsylvania	0.852*	0.361
Rhode Island	0.521	-0.041
South Carolina	-0.225	0.094
South Dakota	0.286	-0.982*
Tennessee	-0.462	-1.027*
Texas	0.585	-0.080
Utah	-0.399	-0.321
Vermont	0.409	-0.169
Washington	-0.162	-0.498*
West Virginia	-1.033*	0.397
Wisconsin	-0.047	-0.439
Wyoming	-0.450	-0.266

[a] UTD is an indicator variable for whether the child is up-to-date in vaccine status, see chapter 3.

[b] Significance level *, **, and *** represent 1%, 5% and 10% significance level.