UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Essays in Applied Health Economics

Autor:
María Dolores de la Mata

Director:
Matilde Cardoso Marta Pinto Machado

Departamento de Economía

Getafe, Septiembre 2011

TESIS DOCTORAL

Essays in Applied Health Economics

Autor: María Dolores de la Mata

Director: Cardoso Marta Pinto Machado

Firma del Tribunal Calificador:

Firma

Presidente: Pedro Pita Barros

Vocal: Manuel Fernández Bagüés

Vocal: Marcos Vera Hernandez

Vocal: Ernesto Villanueva López

Secretario: Jesús María Carro Prieto

Calificación:

Getafe, de de

A mi Familia

Acknowledgements

Es díficil resumir en unas pocas palabras el enorme agradecimiento con todos los que con su ayuda, apoyo, aliento, compañía y cariño me dieron los empujoncitos necesarios para llegar hasta el final de esta etapa. Si bien es fruto del esfuerzo personal, no hubiera sido posible sin cada una de estas cosas. A pesar que quisiera que esta fuera la parte mejor escrita de la tesis, seguramente me faltarán palabras o serán insuficientes. Quedando avisados, entonces, que todo lo dicho acá representa sólo una pequeña porción de lo que me gustaría transmitirles, procedo a mis agradecimientos:

En primer lugar, quiero agradecer especialmente a mi directora, Matilde Machado. Para ella, infinitas gracias por su imprescindible guía y apoyo, por estar siempre atenta a brindarme su ayuda tanto en lo académico como en lo personal. Gracias por tu enorme paciencia para guiarme en los primeros pasos de mi investigación. Por tu confianza en mi trabajo y por transmitirme constantemente motivación y optimismo a lo largo de toda la tesis, combustible escencial para seguir adelante en los momentos más complicados. Por brindarme tu tiempo para mis consultas, para escucharme, para darme sugerencias, sin importar hora y lugar. Y sobre todo, por hacer todas estas cosas con tanta generosidad, y siempre con el toque infaltable de tu entusiasmo y calidez humana.

Un agradecimiento también especial a Joaquín, con quien compartí codo a codo parte importante del doctorado, porque su apoyo y amor fueron fundamentales para llegar hasta acá entera. La lista es muuuy larga, pero en resumidas palabras simplemente gracias Joaco por ser como sos, por transmitirme tus energías, tu practicidad, y tu alegría aún en los momentos más complicados. Por no dudar un segundo en ayudarme cuando lo necesito. No se cómo hubiera llegado hasta acá sin tu compañía. Con vos el camino se hizo tremendamente más fácil y divertido.

A Lucila, amiga incondicional, cuasi-hermana y coautora, por ser también una fuente de continuo apoyo. La lista es muy larga también, y si fuera proporcional a las horas compartidas, no terminaría más. Pero quedándome con lo esencial, gracias Lu por todo lo que aprendí y seguiré aprendiendo con vos, por hacerme reir tanto con tus ocurrencias, por ayudarme cada vez que me ahogo en un vaso de agua, por compartir tu increíble inteligencia con nuevas ideas y sugerencias, y también por tu increíble subjetividad para evaluar positivamente todo lo que yo haga, siempre con la intensión de darme un empujoncito para seguir adelante.

A mis compañeros de doctorado, que con su compañía cotidiana hicieron que los días en la oficina, y fuera de ella, fueran mucho más agradables. Especialmente, gracias Eva, por estar

siempre dispuesta a brindar tu ayuda, por el intercambio de ideas, por la compañía. Sobre todo, gracias por tu buen humor, por tolerar estoicamente a tus vecinas de despacho con el "cotorrear" incesante, por entender finalmente nuestro hablar raro "che", y por hacer tu mejor esfuerzo por transmitirnos toda la españolidad a tu alcance ante nuestras dudas de si no eras en realidad una guiri. A Zoe, por compartir el deporte y por tener siempre una visión distinta de las cosas para hacernos pensar. A Nieves, por las horas compartidas en nuestro "three musketeers' paper", para mí una de las mejores experiencias del doctorado. A Dani, porque a pesar de que jugás en otra liga, en la cual −al menos para nosotros− sos ídolo total, siempre estás dispuesto a escuchar nuestras ideas, discutir nuestros papers y responder pacientemente a todas nuestras preguntas. A Lian, por ser fiel compañero tenístico y vecino de oficina ejemplar. A Ale, Gastón, Luis, Luifa, Juan, Román, Georgi, Heicko, Ming, Efi, Ramiro, Rodo, Joel, Carlos, Doruk, Igor, Vane y Joaquín por las charlas, mates, asados, tenis, paddles, SOSs, viajes y salidas compartidas todos estos años.

A aquellos amigos que tuve la suerte de tenerlos cerca en España. A Pau y Caro, gracias por el año "fantástico" en Getafe, por su alegría, por permitirme acumular anécdotas únicas para reirme por varios años más. A Leo, no puedo más que agradecerte por ser tan bueno y tolerante. A grandes amigos que se quedaron en Córdoba, Malu, Lu, Sole y Oli porque con ellos, junto con Lu y Pau, comencé a experimentar esta extraña locura de pasar millones de horas hablando de economía. Gracias chicos sobre todo por compartir esos años tan lindos de mi vida acompañada de su acentito cordobés.

A todas las personas que me enseñaron y guiaron en mi formación. A mis profesores de la Universidad Nacional de Córdoba y profesores de la UC3M. Un agradecimiento especial para Alberto Figueras, Marcelo Capello, Julio Cáceres, Eugenio Giolito y Matthias Kredler.

A muchas otras personas que directa e indirectamente me han apoyado y acompañado en estos años. A la "comunidad argentina" de Getafe que se va renovando año a año, por todas las reuniones que ayudaron a no sentirme tan lejos. A la familia Berniell, siempre atenta a mis progresos. Gracias Ine, Lau y Jorge por tener siempre abiertas las puertas de su casa. A la familia Coleff, que a pesar de ser tantos, todavía tienen un lugarcito para acordarse de mi. A Marco, por sus enseñanzas. A mis amigas de toda la vida de Santa Rosa, por su amistad incondicional. A Bea y Luis, por todo lo que me eseñaron de su país. A Maxi y Cele, por la compañía los primeros años en España.

Por último, pero no menos importante, quiero agradecer a mi familia. A ellos va dedicada esta tesis. Porque ellos me dan el apoyo, la motivación y el cariño incondicional e imposible de cuantificar, que me ayuda día a día a seguir adelante. A mis padres, Oli y Artur, porque son un gran ejemplo para mí, con sus virtudes y defectos. Porque a base de un increíble, y poco común, esfuerzo lograron cosas que no dejan de despertar mi profunda admiración.

A mis hermanos Manu y Javi, por estar siempre tan pendientes de mí. No es menos lo que los admiro a los dos. Gracias Manu, Alba y Justi por hacerme siempre un lugarcito en Suiza cuando fuera necesario. Gracias Javi y Chechu por contagiarme su alegría. Gracias también a mis tías, Marta y Marisa, por su enorme generosidad y cariño infinito. Porque son realmente mis segundas madres. A mi abuelita, María Teresa, por su ejemplo de fortaleza y por todo el amor incondicional que una abuela puede brindar. A mis tíos Marce y Juan, mis referentes en España, gracias por toda la protección, el cariño y el sol. A mis primos Caro, Pancho y Nico, y mi tía Rita, porque a la distancia nunca dejó de llegarme su cariño. Y a Dorita, por ser otra madre más.

Resumen en Castellano

Mi trabajo de investigación realizado para la obtención del grado de Doctor en Economía consta de tres capítulos: 1) "The effect of Medicaid on Children's Health: A Regression Discontinuity Approach", 2) "Spillovers of Health Education at School on Parents' Physical Activity" (en coautoría con Lucila Berniell y Nieves Valdés), y 3) "Estimating a dynamic discrete choice model of health prevention decisions. An application to flu vaccination".

En el primer capítulo estudio el efecto causal de Medicaid, un programa de seguros de salud público en los Estados Unidos, sobre la utilización de servicios médicos y sobre la salud de los niños afectados por el programa. Para estimar los efectos causales del programa utilizo la metodología denominada Regression Discontinuity (RD). Esta metodología explota la discontinuidad en la probabilidad de participar en Medicaid generada por la regla de elegibilidad del programa, basada en que el ingreso familiar se inferior a un umbral. Una particularidad de este caso de estudio es que los umbrales de elegibilidad varían según estado, y para cada estado, estos umbrales han variado a lo largo del tiempo. Esta característica del programa permite estimar los efectos del programa para diferentes umbrales de ingreso familiar. Los datos utilizados provienen del Panel Study of Income Dynamics (PSID) y de la base suplementaria Child Development Study (CDS).

La principal contribución del trabajo es, primero, mostrar que Medicaid tiene efectos positivos contemporáneos sobre la utilización de servicios de salud preventivos, pero heterogéneos dependiendo del nivel del umbral. Los efectos son positivos y estadísticamente significativos para niveles bajos del umbral (entre 100% y 185% de la línea de pobreza) y no significativos para umbrales altos (entre 185% y 250% de la línea de pobreza). En segundo lugar, encuentro que Medicaid tiene efectos nulos y en algunos casos negativos sobre la salud de los niños en el mediano plazo. Medicaid incrementa la probabilidad de ser obeso y reduce la probabilidad de estar en excelente estado de salud. Los efectos negativos son persistentes entre 1 y dos años después de ser elegible, y luego desaparecen. Estos efectos negativos están presentes tanto a bajos (100-185%) como a altos (185-250%) niveles del umbral.

En el trabajo se discuten posibles mecanismos a través de los cuales Medicaid puede tener efectos negativos sobre la salud de los niños. Una explicación posible es que Medicaid genera un efecto desplazamiento en la cobertura de seguros privados. Los padres pueden verse incentivados a inscribir a sus hijos en Medicaid en lugar de comprar un seguro privado, ya que esto permite liberar recursos para consumos alternativos. Este cambio de seguros de salud puede implicar una caída en la calidad del servicio médico que los niños pueden acceder, generando un efecto negativo indeseado sobre su salud. Asimismo, los períodos de espera necesarios para acceder efectivamente a Medicaid o efectos de consumo negativos pueden generar también impacto negativo en la salud de los niños.

En el segundo capítulo estudiamos el efecto indirecto de la educación para la salud recibida por los niños en las escuelas, sobre los hábitos de salud de sus padres, en particular en la probabilidad de hacer actividad física. En nuestro estudio mostramos que si bien este tipo de programas de educación para la salud pueden estar destinados a mejorar los conocimientos y hábitos de salud de los niños, también pueden tener un efecto indirecto sobre la población no directamente afectada por esta política, en este caso, los padres. También discutimos posibles mecanismos a través de los cuales esta política puede afectar a los padres.

Para estimar el efecto causal de la educación para la salud recibida por los niños en las escuelas sobre los hábitos de actividad física de sus padres, explotamos como "quasi-experimentos" las reformas que se produjeron en los programas de educación para la salud en los Estados Unidos, a nivel estadual, entre los años 1999 y 2005. La estrategia de identificación es una metodología de "triples diferencias" permitiendo distintos tipos de tratamientos, dado que las reformas no fueron homogéneas entre estados. Las bases de datos utilizadas son el Panel Study of Income Dynamics (PSID) y las bases de datos sobre programas de educación para la salud a nivel estadual en EEUU, provenientes de la National Association of State Boards of Education (NASBE) Health Policy Database y de la School Health Policies and Programs Study (SHPPS).

El principal aporte de este trabajo es mostrar que la introducción de educación para la salud en las escuelas primarias tuvo un efecto indirecto positivo y significativo sobre la probabilidad de que los padres realicen actividad física. La probabilidad de hacer actividad física fue 20 puntos porcentuales mayor para aquellos padres cuyos hijos estuvieron afectados por la política, respecto de los padres con similares características pero que cuyos hijos no estuvieron afectados por la política. Los efectos son mayores para los padres con menor educación y menor ingreso. Los efectos solo se encuentran sobre los padres pero no sobre las madres. En el trabajo se proponen dos mecanismos a través de los cuales la educación para la salud en las escuelas puede tener un efecto indirecto sobre los hombres y que a su vez este efecto sea mayor sobre los padres con menor nivel socioeconómico. El primer mecanismo explica por qué los padres y no las madres pueden verse afectados, y está relacionado con decisión óptima de complementar la educación recibida por los niños en las escuelas con cambios de hábitos dentro de la familia. Cuando los padres saben que sus hijos reciben educación para la salud en la escuela, éstos intentan complementarla cambiando ellos mismos sus hábitos relacionados con la salud, de modo de educar a sus hijos con el ejemplo. Dado que existe una especialización por género dentro del hogar en las tareas relacionadas al cuidado de los niños y, en particular, dado que los hombres están especializados en las actividades recreacionales, se espera que el efecto de los programas de educación para la salud en la actividad física de los padres recaiga sobre los hombres y no sobre las mujeres. El segundo mecanismo explica por qué los padres con menor nivel socioeconómico se ven más afectados, y está relacionado con la transmisión de información sobre hábitos saludables

de los niños a sus padres. Un menor nivel socioeconómico está relacionado con un menor conocimiento sobre el cuidado de la salud, por lo tanto los padres con menores conocimientos deberían ser los más afectados con el arribo de nueva información.

En el tercer capítulo realizo una investigación empírica de los determinantes socioeconómicos de las decisiones de vacunación contra la gripe. La evidencia empírica muestra que las decisiones de vacunación tienden a ser muy persistentes entre la población mayor a 65 años. A pesar de esta persistencia, la probabilidad de vacunación también aumenta con la edad y tiende a ser mayor entre los individuos en peor estado de salud. Para estudiar el comportamiento preventivo de los individuos, primero formulo un modelo estilizado de decisiones de prevención primaria a lo largo del ciclo de vida. El modelo está basado en un modelo de capital humano e intenta capturar la importancia de los aspectos dinámicos que motivan las decisiones de prevención. El análisis empírico consiste en estimar la forma reducida de la demanda de prevención que se deriva del modelo teórico utilizando modelos probit dinámicos con datos de panel. Esta metodología permite separar el efecto que tiene la experiencia previa con la vacuna, la heterogeneidad inobservada individual, los factores de riesgo y otras características personales observables sobre la persistencia en las decisiones de vacunación observadas. También analizo cómo los incentivos a vacunarse cambian a lo largo del ciclo de vida. Los datos utilizados corresponden a la población mayor en Estados Unidos y provienen de la Medicare Current Beneficiary Survey (años 2001-2004).

El aporte principal del trabajo es mostrar que tanto la experiencia previa con la vacuna, como la heterogeneidad inobservada individual, los factores de riesgo y otras características personales observables explican parte de la persistencia en las decisiones individuales. En el trabajo se muestra la importancia de controlar por la persistencia en hábitos que genera la experiencia previa sobre las decisiones actuales de prevención. Al tener esto en cuenta, la importancia de los factores de riesgo se reduce sustancialmente. Los resultados también muestran que los incentivos a vacunarse aumentan con la edad a tasa decreciente. Los resultados son consistentes con las predicciones del modelo teórico propuesto. Se discuten finalmente las implicaciones de estos resultados para las políticas públicas que intentan incrementar la cobertura de la vacunación contra la gripe en la población mayor.

Contents

Chapter 1

The Effect of Medicaid on Children's Health: A Regression Discontinuity Approach

Abstract

In this paper I estimate the impact of Medicaid on children's health care utilization and their subsequent health outcomes. I estimate the causal effects using a Regression Discontinuity (RD) design. I exploit the discontinuity generated by Medicaid's eligibility rule, based on family income, on program participation rates. In contrast with a standard regression discontinuity approach, here there are multiple eligibility thresholds that vary across states. This feature allows me to estimate heterogeneous effects of the program at different income thresholds. Using data from the Panel Study of Income Dynamics (PSID) and its Child Development Study (CDS) supplement, first, I find that Medicaid increases the use of preventive medical care. The effect is heterogeneous, and it is positive and significant only at low thresholds −between 100% and 185% of the poverty line. Second, I find that Medicaid has a null or a even negative impact on health outcomes in the medium run. It increases the probability of being obese and it reduces the probability of being in excellent health. The negative effects are persistent between 1 and 2 years after being eligible, and them they vanish, and they appear both at low (under 185%) and high (between 185 and 250%) eligibility threholds levels. One likely explanation for the observed negative effects is that Medicaid induces families to drop private health insurance and, through different channels (e.g., health care quality reductions, waiting periods, negative consumption effects) it generates a negative impact on children's health.

1.1 Introduction

There is strong evidence showing a positive relationship between parental socioeconomic status and children's health, leading to health inequalities in early childhood. To the extent that poor health affects the formation of human capital, health may play a key role in the intergenerational transmission of socioeconomic inequalities (Currie, 2009; Almond and Currie, 2010). Currie (2009) suggests that children's health inequalities may be partially explained by disparities in the access to health care services. The provision of public health insurance coverage to children in low income families facilitates the access to medical care and, therefore, may help to weaken the link between socioeconomic status and health.

The US does not have a universal health care system which makes family income an important factor determining access to health care. Public health insurance programs in the US are designed to improve the access to medical care for low income individuals. Medicaid is a means-tested program and entitles those meeting the required conditions to have public health insurance coverage (Kaiser Commission on Medicaid and the Uninsured, 2010). Medicaid is the largest source of insurance coverage for children in the US, covering about 30% of all children and 59% of low income children.[1]

In this paper, I address three questions. First, I study whether Medicaid contributes to enhance children's utilization of health care services, and, more important, whether it contributes to improve their health outcomes. Second, I analyze whether Medicaid has lagged effects over health. Since health is a stock, the effects of insurance coverage may not be visible immediately but with some lag. Finally, I investigate whether the provision of free health insurance to relatively high income families can have some unintended negative effect on their children's health as a consequence of crowding out private insurance coverage. The provision of a public free health insurance to children in certain ranges of family income may compete with private insurance and induce some of these families to drop the private alternative. If switching from the private to the public occurs then this could have a negative effect on children's health, as long as the switch implies a reduction in the quality of health care. Also, waiting periods until effectively accessing Medicaid may have unintended effects on health. Consumption effects, due to the resources freed when dropping the private insurance, may have negative effects as well, for instance, through an increase in the consumption of "junk" food.

I exploit the particular characteristics of Medicaid's eligibility rules to identify the causal effects of the program on children's outcomes. A child is eligible to receive Medicaid coverage if his family income, as a percentage of the federal poverty line, is below a given threshold.

[1] Low income children are those with family income below 200% of the federal poverty line. Source: Urban Institute and Kaiser Commission on Medicaid and the Uninsured estimates based on the Census Bureau's March 2009 and 2010 Current Population Survey (CPS: Annual Social and Economic Supplements). `http://www.statehealthfacts.org`.

This rule generates a discontinuity in the enrollment rates of children with family income close to the threshold, which allows me to implement a regression discontinuity (RD) design. The eligibility criteria for the Medicaid program are set at the state level, therefore the income threshold that determines eligibility varies among states and has been changing through time. With multiple thresholds, the effects estimated pooling all thresholds are not restricted to the individuals located around a single income threshold, but they are averages of the effects across the different thresholds (Black, Galdo, and Smith, 2005; Bloom, 2009; Carneiro and Ginja, 2009). The multiplicity of thresholds also allows me to investigate whether the effects of Medicaid are heterogeneous for the different thresholds, hence, at different family income levels.

I use data from the Panel Study of Income Dynamics (PSID) and the Child Development Study (CDS) supplement, which provide rich information about children's health and health care utilization as well as detailed information on socioeconomic characteristics of the family. The PSID data allows tracking of children's Medicaid status at different ages through childhood.

In the first part of the paper (Section 1.5) I test the internal validity of the RD design by performing a number of checks which support the RD local assumption stating that eligibility is randomly assigned in the neighborhood of the thresholds. First, I show there is no evidence that families have perfect control over their income so that their children just qualify for Medicaid. Second, I show that the eligibility rule generates a discontinuity in Medicaid enrollment rates at the threshold. Third, I provide evidence that the discontinuity in participation rates at the threshold is not generated by discontinuous changes of other individual characteristics.

My results indicate that Medicaid increases the utilization of health care for preventive purposes (measured as whether the child has visited a doctor at least once in the last 12 months for a routine health check-up) in the same period in which a child is eligible for Medicaid coverage. Allowing for heterogeneous effects across thresholds, I find that Medicaid only has a positive effect on utilization for children at relatively low eligibility thresholds (between 100% and 185% of the poverty line, which I call the *"low" thresholds group* hereafter). Medicaid does not induce higher preventive health care utilization at relatively higher threholds (between 185% and 250% of the poverty line, which I call from now onwards the *"high" threholds group*).[2]

The results also suggest that the short run effects of Medicaid on children's health are null

[2]Note that the effects estimated at higher thresholds are associated with a marginal group of children with higher family income levels. However, in terms of the overall US household income distribution, these targeted families have still low levels of income. For instance, the poverty line for a family of 4 members was $ 21,203 in 2007 and the median household income was $ 50,233. Then, children with family income equal to the "high" thresholds are those with annual family income between $39,226 and $ 53,008, hence only some of them are slightly above the median household income.

in the three outcomes analyzed: probability of being in excellent health, the probability of being obese, and the number of school days missed due to illness. In the medium run −between 1 and 5 years after being eligible for Medicaid coverage− I find that Medicaid only has some negative effects on children's health outcomes. These lagged effects appear after one or two years on the probability of being in excellent health and the probability of being obese. I do not find a clear pattern across threshold groups, and the results indicate that the effects are negative both at low and high thresholds. These negative effects vanish after the third year, and cumulative effects are null afterwards.

I discuss possible mechanisms that may explain explain why Medicaid has some negative effects on children's health outcomes in the medium run, which are consistent with the findings. The negative effects may reflect only a "perception" effect triggered by an increase in the contacts with physicians for preventive checkups, but not a real change of children's health. Parents may be now more aware a health problems their children have. This explanation is more appealing for results obtained at low threshold levels. Also, a "quality" effect, as a consequence of Medicaid crowding out private insurance may explain these results. The quality channel may be more suitable to explain the negative impact of Medicaid at high threshold levels, where the marginal groups are composed of children with higher family income levels. The quality channel explanation states that targeting higher income families with Medicaid may induce a crowding out effect and, although it might not affect health care utilization, it might affect the quality of care a child can have access to. This switch may have undesirable health consequences for children as long as there are health care quality differences between Medicaid and private insurances.[3] In an independent and simultaneous work, Koch (2010) also finds evidence that supports this hypothesis. Finally, I also discussed two alternative mechanisms (waiting times and consumption effects) through which the crowding out effect generated by Medicaid may negatively impact children's health outcomes.

Some previous studies address the question of whether health insurance has a positive effect on children's health. Among those analyzing Medicaid, the results are mixed. For example, Currie and Gruber (1996) find evidence that the expansions in Medicaid eligibility thresholds between 1984 and 1992 increased the utilization of medical care and reduced child mortality. In contrast, Currie, Decker, and Lin (2008) find that expansions in Medicaid eligibility thresholds from 1986 to 2005 had no contemporaneous effects on the health of children between 9 and 17 years old, as reported by their parents. Their estimates, however, suggest that expansions that affected children of ages between 2 and 4 are associated with

[3]Even when Medicaid may induce a crowding out effect at low thresholds, where the marginal affected group has lower levels of family income, it may not have an unintended effect on children's health. The reason is that if insurance quality is a normal good, then this group is more likely to buy, in the absence of Medicaid, low quality private insurances. Hence, at low thresholds, switching into Medicaid is more likely to imply an increase in the quality of care.

better health by the time they are 9-17 years old. Koch (2010) finds a contemporaneous negative impact of Medicaid on children's health and he associates it to a decrease in health insurance quality. He finds that Medicaid reduces the probability that the child has a usual source of care and it increases the child's BMI.

There is also an extensive literature studying the extent to which Medicaid expansions have led eligible families to switch from the private to the public health insurance (Cutler and Gruber, 1996; Lo Sasso and Buchmueller, 2004; Card and Shore-Sheppard, 2004; Ham and Shore-Sheppard, 2005; Gruber and Simon, 2007; Koch, 2010)). None of these papers, except Koch (2010), addresses the consequences of this "crowding-out" effect on children's health.

This paper contributes to the literature in several ways. First, I analyze both the contemporaneous and the lagged effects of Medicaid on different measures of health. The paper by Currie, Decker, and Lin (2008) is among the first to attempt estimating these lagged effects. However, in the cross sectional datasets they use, they must impute the family income and the state of residence of the child, since these variables are not observed during childhood. In contrast, I exploit the panel dimension of PSID data to match past eligibility with current health outcomes. Second, the identification strategy I propose allows for the estimation of Medicaid effects that vary across different levels of income. This identification is similar to Koch (2010), but extends the analysis to the medium run effects. Finally, I propose possible explanations for the existence of persistent negative effects of Medicaid, suggesting that the "crowding-out" effect the public insurance generates may have a cost in terms of children's health.

The remainder of the paper is organized as follows: Section 1.2 describes the Medicaid program; Section 1.3 presents the empirical strategy; Section 3.3 describes the data; Section 1.5 validates the regression discontinuity strategy; Section 3.5 presents and discusses the results; and Section 3.6 concludes.

1.2 Medicaid Program

The Medicaid program was introduced in the late 1960s as a health insurance component for state cash welfare programs targeting low-income single female head families. Medicaid is jointly financed by the federal government and the states. The federal government matches state spending on Medicaid.[4] The program is administered by the states and each state sets its own guidelines regarding eligibility and services, but subject to federal rules requiring minimum levels of coverage and services.

[4]The federal share of Medicaid spending is determined by the Federal Medical Assistance Percentage (FMAP), which varies by state based on state per capita income relative to national average (Kaiser Commission on Medicaid and the Uninsured, 2010).

Medicaid eligibility for children was in its origins tied to the participation in the Aid for Families with Dependent Children (AFDC) program. Since the mid 1980s the linkage between AFDC coverage and eligibility for Medicaid has been gradually weakened, by eliminating the family structure requirements for young children and by allowing states to increase the income thresholds that determine eligibility (Currie and Gruber, 1996). The increase in the thresholds was first a state option, but later minimum levels of coverage were imposed by federal mandates. By April 1990, states were required to offer coverage to all children under 6 years old in families with income up to 133% of the poverty line and, starting in July 1991, they were required to provide coverage to all children under age 19, who were born after September 1983 and lived in households with incomes below 100% of the poverty line. As a result, by the mid-1990s, most children in the US living in households with incomes below 100% of the poverty line, and all young children living in households with incomes below 133% of the poverty line were eligible for Medicaid.

In practice, most states opted to raise the income thresholds beyond 133% of the poverty level and some did further increases using own state funds. States also set different threshold levels for different age groups. In 1997, the Medicaid program for children was augmented by the Children's Health Insurance Program (CHIP), which provided extra funds to expand eligibility for children beyond the existing limits of the Medicaid program. The CHIP program was implemented either by expanding the Medicaid program, or designing a new program, with features that mimic private health insurance (Gruber and Simon, 2007). Some states set thresholds up to 4 times the poverty line.

State Medicaid programs must cover mandatory services specified in federal law in order to receive federal matching funds. Medicaid covers a very comprehensive set of benefits and services for children under 21, defined by the pediatric Medicaid benefit also known as Early and Periodic Screening, Diagnostic, and Treatment (EPSDT) (Kaiser Commission on Medicaid and the Uninsured, 2010). The type of services that Medicaid must cover for children according to the federal rules include screening, preventive, and early detection services.[5] Health care must be made available to correct or ameliorate defects and physical and mental illnesses or conditions discovered by the screening services. Children also have access to physician and hospital services (impatient and outpatient). These services are provided with little or no copayment required (Gruber and Simon, 2007).[6] In terms of the package of services covered, Medicaid tends to be more generous than many private

[5]Screening services include all the following services: comprehensive health and developmental history, immunizations, laboratory tests, lead toxicity screening, vision services, dental services, and hearing services.

[6]Copayments for some services were allowed to be higher for those above 150% of the poverty line since 2005. Cost-sharing for preventive care is prohibited for children. Premiums were prohibited for children until 2005 and remain prohibited for children under 150% of the poverty line. However, for those above 150% the poverty line, premiums and cost sharing cannot exceed 20% of the cost of the service. Additionally, total premiums and copayments cannot exceed 5% of family income for any family (Kaiser Commission on Medicaid and the Uninsured, 2010).

insurance plans.

Medicaid buys services primarily in the private health care sector. States pay health care providers on behalf of the Medicaid beneficiaries. States may purchase services on a fee-for-service basis or by paying premiums to managed care organizations (Kaiser Commission on Medicaid and the Uninsured, 2010). States also determine the rules to reimburse health care providers. In most cases, Medicaid's reimbursement is lower than the obtained from private insurance, which may induce some physicians to reject Medicaid patients or to lower the quality of the service provided.[7]

1.3 Empirical Research Design

1.3.1 Contemporaneous Effects

The main objective is to estimate a simple model of the causal effect of Medicaid coverage on children's health care utilization and health outcomes

$$y_{it} = \alpha + \beta M_{it} + u_{it}, \tag{1.1}$$

where y_{it} is child i's outcome (utilization or health) in period t and M_{it} indicates whether the child had Medicaid coverage that same period. A simple OLS regression of equation (1.1) would yield a biased estimate. Medicaid coverage is an endogenous variable, because the access to this type of coverage is correlated with family income. Even after controlling for family income, selection problems may still be present because Medicaid enrollment is not mandatory. Among eligibles, the decision to take Medicaid may be correlated with other unobserved characteristics that are correlated with the outcomes.

In order to identify the effect of interest, I exploit the rule of assignment into Medicaid that allows me to implement a Regression Discontinuity (RD) design. The RD design is a quasi-experimental design with the defining characteristic that the probability of receiving the treatment changes discontinuously as a function of the variable that determines eligibility, called the assignment or forcing variable (Hahn, Todd, and Van der Klaauw, 2001).[8]

The intuition behind the RD is the following. Assuming that the eligibility threshold is exogenously given and families have imperfect control over their income, the eligibility status of a child with family income in the neighborhood of the threshold is randomly

[7] For example, Decker (2007) finds that higher Medicaid fees increase the number of private physicians, especially in medical and surgical specialties, who see Medicaid patients. She also finds that higher fees also lead to visit times with physicians that are more comparable to visit times with private pay patients. Another paper by Cunningham and O'Malley (2009) finds that not only reimbursement fees matters, but also delays in reimbursement. They find evidence that Medicaid reimbursement time affects physicians' willingness to accept Medicaid patients.

[8] For a comprehensive discussion of the RD design and its application in economics see Imbens and Lemieux (2008), van der Klaauw (2002), and Lee and Lemieux (2010)

assigned, i.e., the rule generates a "local" randomized experiment. Making the additional assumption that in the absence of the treatment the outcome is a smooth function of income, the causal effect of Medicaid eligibility can be identified by comparing the average outcome of children just below the income threshold ("treatment group") with that of children just above it ("control group"). Any difference observed between these two groups can be attributed to the availability of treatment for treatment group members. Since enrollment in Medicaid is not mandatory −i.e., the coverage indicator, M_i, is not necessary equal to an indicator of eligibility status, Eli_i, which takes the value one if the child is eligible for Medicaid− comparing outcomes of eligible and non eligible individuals close to the threshold identifies the average effect of assignment into treatment or the *intention to treat effect* (ITT) at the threshold.[9]

The ITT effect can be significantly lower in absolute value than the effect the program has on those who are actually covered by Medicaid. Under the assumptions that the probability of having Medicaid coverage as a function of income is discontinuous at the threshold and that, in the absence of the treatment, the association between the outcome variable and income is smooth, the parameter β can be estimated using the eligibility indicator Eli_i −which is randomly assigned in the neighborhood of the threshold− as an instrument for Medicaid coverage. This is called a "fuzzy" RD design (Hahn, Todd, and Van der Klaauw, 2001; Imbens and Lemieux, 2008).[10]

Ideally, to identify the causal effect it would be sufficient to compare outcomes of individuals above and below the threshold, in a very narrow interval around it. In practice, however, this is sometimes not possible because only few observations close to the threshold are available in the dataset. To overcome this problem, I implement a parametric RD specification, that controls for a flexible function of the family income −assignment variable. I estimate β by 2SLS, where I instrument the treatment dummy, M_i, with the eligibility status, Eli_i.

The two equation system is given by

$$y_{it} = \alpha + \beta M_{it} + k_{2g}(z_{it}) + k_{2g}(z_{it}) \times Eli_{it} + u_{it}, \tag{1.3}$$

$$M_{it} = \pi_0 + \pi_1 Eli_{it} + k_{1g}(z_{it}) + k_{1g}(z_{it}) \times Eli_{it} + v_{it}, \tag{1.4}$$

[9]For instance, studies such as Currie and Gruber (1996) and Currie, Decker, and Lin (2008), although using different identification strategies than in this paper, identify the intent to treat effects of Medicaid on children who where newly eligible to receive Medicaid benefits with the Medicaid expansion.

[10]As shown by Hahn, Todd, and Van der Klaauw (2001), the treatment effect can also be recovered by dividing the "jump" in the relationship between the outcome and eligibility −the ITT at the threshold− by the fraction of individuals induced to take Medicaid at the threshold

$$\beta = \frac{\lim_{z \to z_0^-} E[y_i|z_i = z] - \lim_{z \to z_0^+} E[y_i|z_i = z]}{\lim_{z \to z_0^-} E[M_i|z_i = z] - \lim_{z \to z_0^+} E[M_i|z_i = z]}, \tag{1.2}$$

where z_i is the family income and z_0 is the eligibility threshold.

Hahn, Todd, and Van der Klaauw (2001) were the first to show the connection between how the treatment effect is defined in the fuzzy RD design and the estimation of the treatment effect in an instrumental variables setting, when the instrument is a binary variable.

where $Eli_{it} = \mathbf{1}\{\frac{z_{it}}{PL_t} \leq T_t\}$, is a dummy variable that takes the value one if the child is eligible for Medicaid, i.e., when family income (z_{it}), as a percentage of the poverty line (PL_t), is below the eligibility threshold (T_t); $k_{1g}(.)$ and $k_{1g}(.)$ are polynomials of order g of family income which are allowed to be different at each side of the threshold, and u_{it} and v_{it} are unobserved error components. Given that the poverty line varies with family size,[11] in all specifications I control for this variable. Additionally, some states set different thresholds for different children's age groups, hence, I also control for child's age. The periods for which I observe the outcomes are t=1997, 2002, 2007, as I explain in Section 3.3. Since the model is exactly identified, 2SLS estimates of β are numerically identical to the ratio of the reduced form coefficients θ/π_1, provided the same order of polynomial is used for $k_1(.)$ and $k_2(.)$ (Lee and Lemieux, 2010).

The parametric specification in equation (1.3) allows to retain observations that are not necessarily close to threshold. The polynomial function of income controls for variation in the outcome and participation coming from income differences far from the threshold. Hence, β captures differences in the outcome variable for individuals just at the threshold. To check the robustness of the results, as suggested by Lee and Lemieux (2010), I run the regressions narrowing the width of the interval in the neighborhood of the threshold and I control for different orders of polynomials, selecting the optimal order for each bandwidth according to the Akaike information criterion (AIC) of model selection.

Hahn, Todd, and Van der Klaauw (2001) were the first to suggest estimating the treatment effect in the fuzzy RD setting using two-stage least-squares (2SLS). Furthermore, they also point out that the estimate of β can be interpreted as a *Local Average Treatment Effect* (LATE) at the threshold under the same assumptions as in Imbens and Angrist (1994). Under these assumptions, the LATE is defined as the average effect of treatment on the population of *"compliers"*, those eligible individuals at the threshold who receive the treatment if and only if they are assigned to it.

I also estimate the reduced-form equation that recovers the IIT effects

$$y_{it} = \alpha + \theta Eli_{it} + f_g(z_{it}) + f_g(z_{it}) \times Eli_{it} + u_{it}, \tag{1.5}$$

where $f_g(z_{it})$ is a polynomial of order g of income. The parameter θ captures the ITT effect at the threshold, and, given that there is not perfect compliance, this parameter is always a lower bound of β.

Medicaid is a state administered program in which each state sets its own eligibility threshold, hence there are multiple thresholds at a given point in time. Therefore, to implement the RD design some clarifications are required. First, the RD assumptions should hold

[11] For example, the poverty line in 2007 was \$16,530 for a family of 3 members and \$21,203 for a family of 4 members.

for each state in each year. In principle, I could estimate the model in equation (1.3) or (1.5) for each separate state and year. However, due to sample size restrictions, throughout the paper I either pool all the thresholds (pooling states and years, which I call the "full sample") or I divide the sample in two subgroups of thresholds ("high" and "low" threshold groups). Second, I need to make family income comparable in terms of the distance from their respective eligibility threshold, so I rescale it subtracting the respective threshold. Third, the pooled analysis imposes a common treatment effect β or a common intention to treat effect θ at each threshold (i.e., the eligibility indicator in equations (1.3), (1.4), and (1.5) takes the value 1 if family income is below the threshold, but it does not distinguishes which threshold it is). Hence, the estimated effect is a pooled treatment effect or a pooled intention to treat effect, respectively, as in Carneiro and Ginja (2009) and Listing (2010). I also allow for Medicaid effects to differ between groups of thresholds. I define as "low" thresholds all those thresholds that are between 100% and 185% the poverty line and "high" threhsols all those threhsolds between 185% and 250%.[12] I assume a common effect of Medicaid across all thresholds that are classified as "low" thresholds and a common effect across all thresholds classified as "high" thresholds. I allow these two effects to be different and test wehther the difference is statistically significant.[13] When I allow for heterogenous effects across groups, the estimated models are

$$\begin{aligned} y_{it} = \quad & \alpha_0 + \beta_0 M_{0it} + \beta_1 M_{1it} + \alpha_1 T_{1it} + k_g(z_{it}) \\ & + k_g(z_{it}) \times T_{1it} + k_g(z_{it}) \times Eli_{1it} + k_g(z_{it}) \times Eli_{1it} \times T_{1it} + u_{it}, \end{aligned} \tag{1.6}$$

and

$$\begin{aligned} y_{it} = \quad & \alpha_0 + \theta_0 Eli_{0it} + \theta_1 Eli_{1it} + \alpha_1 T_{1it} + f_g(z_{it}) \\ & + f_g(z_{it}) \times T_{1it} + f_g(z_{it}) \times Eli_{1it} + f_g(z_{it}) \times T_{1it} \times Eli_{1it} + u_{it}, \end{aligned} \tag{1.7}$$

where $T_{j,it}$ is an indicator that takes the value one if child i lives in period t in a state where the eligibility threshold is T_j% of the poverty line, and $Eli_{j,it} = Eli_{it} \times T_{j,it}$, is an indicator that takes the value one if the child is eligible for Medicaid and lives in a state where the eligibility threshold is T_j%. I consider two categories of T: thresholds lower than 185% of the poverty line (baseline category, T_0), thresholds between 185% and 250% of the poverty

[12] Some states set thresholds above 250% the poverty line. I do not consider them because, as I discuss in Section 5.2, there is not a visible discontinuity in the probability of participating in Medicaid at these higer threholds.

[13] Note that, in general, the effects of Medicaid at different threshold levels will be estimated with different states −because not all states have the same threshold in a given moment of time, except those states that set different eligibility threhold for different age groups. To be able to compare the effects of the low and high threhold groups, in the sense that the only difference across groups is the income level but not other state specific characteristic, we require no correlation between the state threshold level and state characteristics, which are correlated with the potential effect the state Medicaid program may have on children's outcomes.

line (T_1). I instrument medicaid coverage with the eligibility indicators, $Eli_{0,it}$ and $Eli_{1,it}$.

1.3.2 Lagged Cumulative Effects

In order to estimate the medium run causal effects of Medicaid on children's health I also take advantage of the "local" random assignment that the eligibility rule generates in a period $t-\tau$ to estimate the effects that Medicaid has, τ periods later, on period t outcomes using the following specification

$$y_{it} = \alpha + \theta_\tau Eli_{i,t-\tau} + f_g(z_{i,t-\tau}) + f_g(z_{i,t-\tau}) \times Eli_{i,t-\tau} + u_{it}, \tag{1.8}$$

where $Eli_{i,t-\tau}$ is a dummy variable that takes the value one if the child was eligible for Medicaid in period $t-\tau$ and f_g is a polynomial of order g of income in period $t-\tau$, $z_{i,t-\tau}$.[14] The parameter θ_τ does not isolate the direct effect of eligibility in period $t-\tau$ on period t outcomes, because of the possibility of multi-treatment. That is, between periods $t-\tau$ and t a child may have multiple opportunities to be eligible and enrolled in Medicaid. To the extent that period $t-\tau$ eligibility affects posterior participation in Medicaid, then the parameter θ_τ will also capture the indirect effect that subsequent participation may have on health outcomes of period t.

Given the possibility of multi-treatment, the marginal effect of making a child randomly eligible for Medicaid in a period $t-\tau$ on health outcomes in period t reflects a cumulative effect which is the sum of: 1) a direct effect on health outcomes τ years later, if it were possible to prohibit the child from being assigned to treatment in any other subsequent period; 2) an indirect effect on health outcomes through the effects on subsequent participation in the program. The total effect or *medium run ITT effect* of Medicaid eligibility on subsequent health, captured by θ_τ, is the effect of exogenously making a child eligible in a given period, without controlling for the family behavior in subsequent years. Following Cellini, Ferreira, and Rothstein (2010) the ITT parameter is[15]

$$\theta_\tau^{ITT} = \frac{\mathbf{d}y_{it}}{\mathbf{d}Eli_{i,t-\tau}} = \frac{\partial y_{it}}{\partial M_{i,t-\tau}} \times \frac{\partial M_{i,t-\tau}}{\partial Eli_{i,t-\tau}} + \sum_{h=1}^{\tau} \left(\frac{\partial y_{it}}{\partial M_{i,t-\tau+h}} \times \frac{\partial M_{i,t-\tau+h}}{\partial Eli_{i,t-\tau}} \right), \tag{1.9}$$

where $\frac{\partial y_{it}}{\partial M_{i,t-\tau}}$ is the direct effect of Medicaid in period $t-\tau$ under the assumption that the child would not have access to Medicaid in the subsequent years, and $\frac{\partial M_{i,t-\tau+h}}{\partial E_{i,t-\tau}}$ is the effect that eligibility in period $t-\tau$ has on subsequent Medicaid participation. For the medium run analysis I also consider specifications that allow for heterogeneous effects effects by

[14] For the medium run analysis I restrict to estimating the ITT effects given that these effects provide a lower bound of the average treatment effects and the IV estimates tend to be more imprecisely estimated.

[15] The main difference with Cellini, Ferreira, and Rothstein (2010) is that in their paper they have a "sharp" RD design, that is, being eligible is equivalent to receiving the treatment.

eligibility threshold levels.

1.4 Data

The datasets used in the analysis are the Panel Study of Income Dynamics (PSID) and the Child Development Study (CDS) supplement. The CDS is a sample of children who were between 1 and 12 years old by 1997 and it contains information about children's health care utilization and health outcomes, obtained from the children's primary caregiver, as well as characteristics such as age and race of the child. Data for this cohort of children were collected in three waves: 1997, 2002, and 2007. Information on family income, Medicaid coverage, and family characteristics comes from the PSID dataset which can be matched with the CDS. I use the three CDS waves-matched with PSID data as repeated cross-sections, and I restrict the sample to children between 5 and 18 years old in any of the three waves. I keep only children for whom I can keep track of their eligibility and Medicaid status up to 5 years before the outcomes are observed.

I assign the Medicaid eligibility status of each child in the survey on a yearly basis. To impute eligibility I compare the annual family income as a percentage of the poverty line with the corresponding eligibility threshold, that is

$$Eli_{it} = \mathbf{1}\{\frac{\text{income}_{it}}{PL_t(fasize_{it})} \leq T_t(\text{state}_{it}, \text{age}_{it}\}, \tag{1.10}$$

where PL_t is the federal poverty line in period t (a function of the family size), and $T_t(.)$ is the state-age specific threshold in period t. I use the family income and the annualized official poverty threshold provided in the PSID data file for each family.[16] I get the information of state-age-year specific threshold from various reports of the National Governors' Association.

I use three types of outcome variables: one measure of preventive health care utilization; one objective measures of health; and two subjective measures of health. The measure of preventive health care utilization is a variable that indicates whether the child had visited a doctor at least once in the last 12 months for a routine health check-up. This measure is generally used to capture the utilization of medical resources for preventive purposes.[17]

As an objective measure of health I use the Body Mass Index (BMI).[18] A child's weight status is determined based on an age and sex-specific percentile for BMI. A child is classified

[16] See Grieger, Danziger, and Schoeni (2009) for further details on the measures of family income and poverty thresholds in PSID. All income measures are expressed in 2000 US dollars.

[17] Other measures of health care utilization, such as the number of hospitalizations, may confound access and morbidity, as pointed out by Currie and Gruber (1996). An absence of a doctor visits for a regular check-up, however, better reflects an "access" problem.

[18] Although it is not completely "objective", since during the interview, the primary caregiver reports the weight of the child, and the interviewer measures his or her height.

as obese if her BMI is at or above the 95th percentile of the BMI distribution of children of the same age and sex.[19] Medicaid coverage may facilitate and increase the contact with physicians, which in turn increases the likelihood that children's weight status is monitored. Physicians recommendations about the quality of the diet and the adequate level of physical activity may be critical inputs to improve children's health status.

Additionally, I use two subjective health measures, both reported by the child's caregiver: an indicator of whether the child has an excellent health status and the number of school days the child missed due to illness during the last 12 months. The first measure reflects the caregiver's perception about the child's overall health status.[20] The second measure links child's health status and school attendance, capturing a key aspect of how health may affect her human capital formation. If Medicaid allows to prevent illnesses it might also help to avoid missing school days.

One drawback of measuring the effects of Medicaid on subjective health measures is that these effects may be difficult to interpret. Currie and Gruber (1996) argue that these measures may capture two possible effects. If the public insurance coverage leads individuals to increase the contacts with the medical system, then there could be a "true" effect on child health, resulting in better child's health reports. The increased contacts with physicians, however, may also affect parents' perception about the health of the child. Parents may learn about health conditions the child already had but they were not aware of because they did not contact physicians so frequently before having the public insurance coverage. Also, if targeted children are switching from a private insurance to the public, parent's reports may be sensitive to perceived changes in the quality of health care they have access to with the public insurance instead of the private one.

Columns (1) and (2) of Table 3.2 present descriptive statistics of children's and family main characteristics, for the full sample. I refer as "full" sample as the sample pooling all eligibility thresholds. Here, I consider all children whose annual family income is within a distance of $\pm$ 50 thousand dollars in period t, for t=1997, 2002, and 2007, although for the empirical analysis I restrict to narrower intervals in the neighborhood of the threshold. Columns (3) to (8) present the same descriptives but for three subsamples, defined by the level of Medicaid "generosity" in each state, where the generosity is determined according to the level of the income threshold, as a percentage of the poverty line, that determines eligibility. The first subsample consists of children living in states where the generosity of Medicaid coverage is relatively low (the eligibility thresholds are lower than 185% the

[19]The CDS dataset provides indicators of the child's obesity and overweight status according to this definition, based on the Centers for Disease Control and Prevention (CDC) growth charts. Each of the CDC BMI-for-age gender specific charts contains a series of curved lines indicating specific percentiles. See the CDC Growth Charts for children at: http:\www.cdc.gov\growthcharts.

[20]Parent's report about child's health status can fall in one of four categories: Excellent, Good, Fair, and Poor. In general parents tend to report either Excellent or Good, hence I constructed a dummy variable that takes the value one when excellent is reported and zero otherwise.

poverty line); the second subsample includes children living in states with a middle level of generosity (the eligibility thresholds are set between 185% and 250% the poverty line); and finally, the third subsample consist of children living in states with relatively high levels of generosity (the eligibility thresholds are above 250% the poverty line).[21]

From columns (1) and (2) it is clear that Medicaid eligibles are more disadvantaged than non-eligibles in several dimensions. They have lower family income −by definition of eligible−, they are more likely to be minorities, to live in a female-headed family, and to live with a less educated head of household. They are worse off in terms of health outcomes. However, they are more likely to have visited a doctor for a check up in the last 12 months. A similar pattern emerges if I split the sample according to the different levels of Medicaid's coverage generosity. In the three groups, eligible children are always more disadvantaged in terms of socioeconomic characteristics, they tend to have worse health outcomes, and to use more preventative health are services. The only exception is the states with higher levels of generosity, where utilization is higher for non-eligibles.

Only 54% of eligible children are actually enrolled in Medicaid, although enrollment is heterogeneous depending on family income level.[22] The incentives to enroll in Medicaid decrease with income, as it can be observed by comparing eligible children in states with higher levels of Medicaid generosity. The take-up rate is 64% in states with modest Medicaid coverage generosity, where eligibles' average family income is 11.8 thousand dollars per year. This proportion falls to 53% in states with middle level generosity and where eligibles' average income is 20.9 thousand dollars, and it is even lower (39%) in states with the most generous coverage, where eligibles' average income is 33.0 thousand dollars. The incentives to enroll in Medicaid may decline with income because, as income rises, the family's financial constraint is less binding, which allows them to acquire an alternative source of coverage in private markets.

[21]The choice of the limits of the groups is done in such a way that the 200% eligibility threhold falls in the second group, because of the large number of observation associated to this cuttof. I also tried alternative groups and in general the reults of the paper are not sensitive to this classification. Alternative grouping were: $T \leq 185$, $185 < T < 250$, and $T \geq 250$, and $T < 200$, $T = 200$, and $T > 200$.

[22]Note that among non-eligibles there are individuals with Medicaid coverage. This happens because there may be timing problems in the reports of individuals family income −from which I infer eligibility status− and Medicaid coverage. Also, income fluctuations during the year can make an individual eligible for Medicaid at some point of the year but according to the annual income I they re clasified as non-eligible. Approximately 10% of the non-eligibles in the full sample report having Medicaid, although this percentage rise up to a 20% for the subgroup of individuals just above the eligibility threshold, as will be showed in Section 1.5.

Table 1.1: Descriptive Statistics

	Full sample		Thresholds [100%-185%) PL		Thresholds [185,250]% PL		Thresholds>250% PL	
	Eligible (1)	Non-Eligible (2)	Eligible (3)	Non-Eligible (4)	Eligible (5)	Non-Eligible (6)	Eligible (7)	Non-Eligible (8)
Outcome Measures								
Visited a doctor at least once in last 12 months	0.75	0.68	0.73	0.63	0.77	0.72	0.74	0.75
Obese	0.23	0.20	0.22	0.20	0.23	0.20	0.24	0.18
Overweight	0.16	0.15	0.15	0.15	0.16	0.14	0.15	0.15
Obese + Overweight	0.39	0.35	0.37	0.35	0.39	0.35	0.40	0.33
BMI	21.77	21.06	20.00	19.91	22.43	22.10	22.19	21.54
School days missed due to illness	2.60	2.64	3.23	2.38	2.33	2.88	2.65	2.58
	(5.76)	(4.86)	(7.58)	(3.61)	(5.14)	(5.86)	(3.77)	(3.67)
Health Excellent	0.41	0.52	0.34	0.47	0.43	0.56	0.45	0.60
Insurance Coverage								
Medicaid Coverage	0.54	0.11	0.64	0.13	0.53	0.09	0.39	0.07
Private Insurance	0.30	0.78	0.20	0.76	0.30	0.79	0.53	0.88
Medicaid and Private	0.04	0.02	0.06	0.03	0.04	0.02	0.03	0.02
Family and Child Characteristics								
Family Income (2000 dollars)	19,916.53	52,139.65	11,799.60	43,732.20	20,868.69	56,673.64	33,023.05	78,163.68
	(12,412.04)	(18,517.90)	(7,371.64)	(15,269.43)	(10,788.69)	(16,531.64)	(16,248.02)	(19,740.05)
Income Cutoff (Poverty Line × T)	35,689.21	30,314.07	22,589.13	21,279.56	37,397.48	35,195.15	55,891.68	58,202.88
	(13,298.81)	(12,435.80)	(6,503.69)	(6,058.63)	(9,078.54)	(8,236.56)	(13,876.26)	(13,979.42)
Income threshold as % of poverty line	198.28	172.06	124.17	119.72	203.51	200.80	337.17	329.69
	(64.51)	(59.51)	(23.72)	(23.61)	(15.62)	(15.34)	(46.14)	(41.66)
Metropolitan Area	0.72	0.65	0.77	0.65	0.70	0.64	0.67	0.82
Rural Area	0.13	0.13	0.10	0.12	0.13	0.15	0.15	0.07
Family Size	4.27	4.07	4.29	4.10	4.35	4.04	3.79	4.05
	(1.41)	(1.17)	(1.37)	(1.14)	(1.42)	(1.19)	(1.30)	(1.16)
Education (yrs.) of the Head of the Household	11.58	12.88	11.61	12.70	11.40	12.91	12.46	14.02
	(2.49)	(2.07)	(1.98v	(2.04)	(2.70)	(2.05)	(2.08)	(2.11)
Female Head	0.61	0.26	0.71	0.29	0.59	0.24	0.49	0.14
Child Age	11.03	10.84	9.46	9.54	11.59	11.89	11.48	11.75
	(3.26)	(3.28)	(3.13)	(2.95)	(3.14)	(3.13)	(3.01)	(3.42)
Male	0.52	0.51	0.49	0.50	0.52	0.52	0.55	0.53
Black	0.62	0.40	0.75	0.47	0.60	0.34	0.42	0.28
Birth Weight (kg)	3.19	3.34	3.15	3.34	3.19	3.34	3.24	3.36
	(0.69)	(0.64)	(0.72)	(0.65)	(0.67)	(0.64)	(0.71)	(0.60)
Mother Age at Child's Birth	24.96	26.84	24.80	26.45	24.92	27.04	25.51	28.18
	(5.87)	(5.57)	(5.71)	(5.49)	(5.95)	(5.63)	(5.78)	(5.38)
N	1577	2115	414	972	986	1024	177	119

Notes: Observations are restricted to children in the CDS whose family income is at a distance of ± 50 thousand dollars from the threshold in years 1997, 2002 or 2007. Columns (1) and (2) present descriptive statistics for the full sample. Columns (3) and (4) correspond to the subsample of children living in states where the generosity of Medicaid coverage is relatively low −the eligibility thresholds are lower than 185% of the poverty line; Columns (5) and (6) correspond to the subsample of children living in states with a middle level of generosity −the eligibility thresholds are set between 185% and 250% of the poverty line; Columns (7) and (8) correspond to the subsample of children living in states with relatively high levels of generosity −the eligibility thresholds are above 250% of the poverty line.

1.5 Validity of RD Design: Robustness Analysis

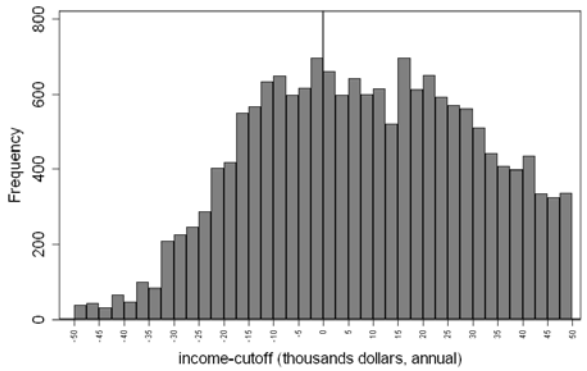
A first step in the analysis involves testing the identification assumptions of the RD to check its internal validity. The empirical strategy is based on the assumption that eligibility to receive Medicaid coverage is as good as randomly assigned in the neighborhood of the income thresholds. This assumption requires families to be unable to manipulate their incomes perfectly well so that they cannot control if their children qualify for Medicaid. Additionally, for the validity of the design, the probability of participating in Medicaid as a function of family income should show a discontinuity at the threshold. Finally, an implication of the "local" randomization is that individuals at either side of the threshold should be similar both in observed and unobserved characteristics.

To check the validity of the RD design I perform the following robustness analysis, as proposed by Imbens and Lemieux (2008) and Lee and Lemieux (2010). First, I inspect the histogram of the family income −the assignment variable− to check whether families have imprecise control over it. A spike to the left of the threshold may indicate that families are manipulating their income to fall below the eligibility threshold. Second, I estimate the participation equation to check whether Medicaid eligibility rule induces a discontinuity at the threshold. Finally, I examine whether baseline covariates (variables that should not be affected by the program as well as individual characteristics not taken into account to determine eligibility) are balanced on either side of the threshold.

1.5.1 Manipulation of the assignment variable

Figure 1.1 presents an histogram with the distribution of family income, pooling all observations for all years in which I observe the family income of children in the CDS (1991-2007). This graph shows the number of observations within bins of 2.5 thousand dollars width. Given that there are multiple thresholds, income is normalized by subtracting the corresponding eligibility threshold. A negative value indicates that income is below the threshold and the child is eligible for Medicaid. An accumulation of observations below the normalized threshold (equal to zero) may be an indication of income manipulation. At first sight families do not seem to be manipulating their income in order to be below the threshold.

Figure 1.1: Distribution of the Family Income (normalized). All thresholds pooled. Years 1991-2007.



McCrary (2008) proposes a simple two-step procedure for testing whether there is a discontinuity in the density of the assignment variable. Figure 1.2 graphically displays the results of the density discontinuity test at the cutoff for different samples. Figure A presents the density estimate for the full sample. The estimate of the log difference in the height of the density function at the threshold is -0.027 (standard error 0.045). The test suggests no discontinuity in the density at the normalized threshold (t-statistic of -0.590).[23]

To check whether the incentives to manipulate family income vary across different eligibility thresholds I perform the test on three subsamples. The first subsample consists of families who reside in states where the eligibility threshold is lower than 185% the poverty line (Panel B of Figure 1.2), the second subsample considers families who reside in states with eligibility thresholds between 185% and 250% the poverty line (Panel C of Figure 1.2), and finally, a the third subsample consists of families residing in states with thresholds above 250% the poverty line (Panel D of Figure 1.2). In all the cases the test fail to reject the null hypothesis of no discontinuity at the threshold.

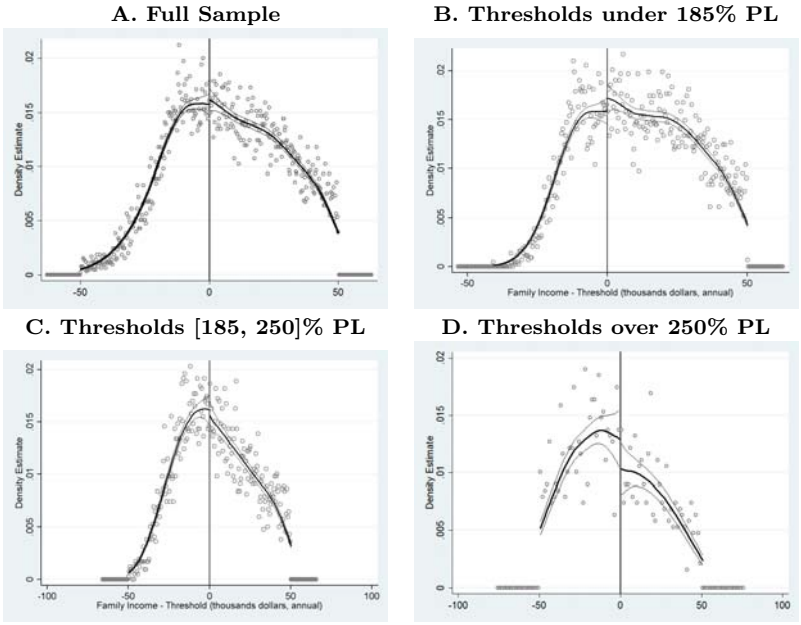
1.5.2 Discontinuity in the probability of participating in Medicaid

As discussed in Section 1.3, despite of imperfect compliance, the fuzzy RD analysis can identify a LATE at the threshold as long as the eligibility rule generates a jump in the participation rate at the threshold.

Figure 1.3 plots the proportion of children who are enrolled in Medicaid over family income for the years 1997, 2002, and 2007. Each dot is the proportion of children with Medicaid coverage within a family income bin of 2 thousand dollars width. The solid lines are predictions from local linear regressions with bandwidth of 5 thousand dollars estimated with the raw data. We can observe that at the threshold −normalized to 0− the probability of participation has a discontinuity of about 15 percentage points.

[23] I also I perform the same exercise on the sample that that pools years 1997, 2002, and 2007. I also reject the null hypothesis of a discontinuity of the density distribution at the threshold.

Figure 1.2: Testing Manipulation of Assignment Variable. Years 1991-2007.



Dots are density with the indicated binsize (in thousands dollars). Solid lines are predictions from local linear regressions using triangle kernel. Standard errors, binsize b, and the bandwidth h are calculated as in McCrary (2008). **Full sample:** h=12.82 b=0.288 **Under 185:** h=12.24 b=0.349. **185-250:** b=0.481 h=15.865. **Over 250:** h=25.59 b=1.388

Table 1.2 reports the results of the parametric estimation of the participation equation specified as

$$M_{it} = \pi_0 + \pi_1 Eli_{it} + k_{1g}(z_{it}) + k_{1g}(z_{it}) \times Eli_{it} + u_{it}, \qquad t = 1997, 2002, 2007, \quad (1.11)$$

where Eli_{it} is the eligibility indicator in period t, M_{it} is Medicaid enrollment status in the same period, and $k_{1g}(.)$ is a polynomial of order g of family income, z_i.

The table presents the estimated jump in the full sample (pooling all thresholds), and in three separate subsamples for different groups of threshold levels. Each column of this table shows the estimates of the same model but considering windows of different widths around the threshold.[24] The results indicate that making a child with family income equal to the threshold eligible for Medicaid increases the probability of enrollment by between 12 and 16 percentage points in the full sample, depending on the width of the interval around the threshold.[25]

The results show that the discontinuity in Medicaid participation is larger in states with

[24] I select the order of the polynomial according to the Akaike information criterion (AIC) of model selection, $AIC = N \ln(\hat{\sigma}^2) + 2p.$, where $\hat{\sigma}^2$ is the mean squared error of the regression, and p is the number of parameters in the regression model. The AIC favors in most of the cases polynomials of order 1 or 2.

[25] Excluding the case for a bandwidth of 30 and polynomial of order 1, that seem to overestimate the jump at the threshold.

Figure 1.3: Medicaid Participation. Full sample. Years 1997, 2002 and 2007.

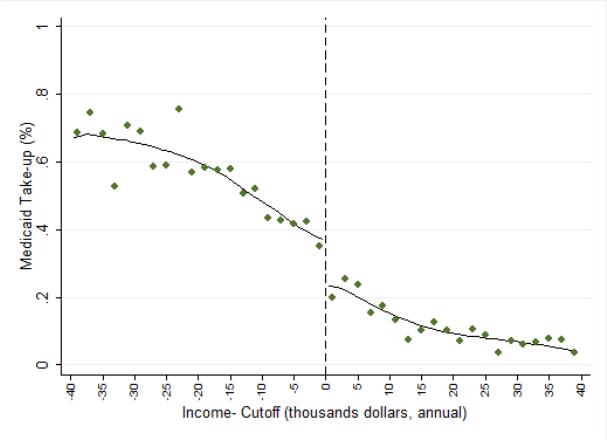


Table 1.2: Participation Equation. "Jump" at the threshold. Years 1997, 2002, 2007.

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
One						
All Thresholds	0.236***	0.108*	0.096	0.251***	0.178***	0.149**
	(0.056)	(0.057)	(0.070)	(0.052)	(0.059)	(0.068)
Thresholds under 185%	0.275***	0.220***	0.142**	0.248***	0.164***	0.128*
	(0.056)	(0.057)	(0.071)	(0.057)	(0.061)	(0.075)
Thresholds 185-250%	0.236***	0.108*	0.096	0.251***	0.178***	0.149**
	(0.056)	(0.057)	(0.070)	(0.052)	(0.059)	(0.068)
Thresholds over 250%	-0.003	0.059	-0.070	0.039	0.088	-0.016
	(0.088)	(0.108)	(0.125)	(0.083)	(0.108)	(0.113)
Two						
All Thresholds	0.136***	0.122***	0.126***	0.149***	0.163***	0.137**
	(0.036)	(0.039)	(0.047)	(0.041)	(0.046)	(0.058)
Thresholds under 185%	0.176***	0.154**	0.123	0.155**	0.159**	0.150
	(0.063)	(0.069)	(0.083)	(0.073)	(0.079)	(0.106)
Thresholds 185-250%	0.110**	0.144**	0.174**	0.173***	0.174**	0.227**
	(0.050)	(0.066)	(0.073)	(0.061)	(0.079)	(0.094)
Thresholds over 250%	0.031	0.102	0.054	0.088	0.168	0.255*
	(0.088)	(0.108)	(0.101)	(0.109)	(0.120)	(0.131)
Different Polynomial at each side	N	N	N	Y	Y	Y
N	2163	1738	1361	2163	1738	1361

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of the indicated order of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated.

the lower eligibility thresholds, except in the smaller bandwidth considered. In states where $T < 185\%$ the jump is between 12 and 22 percentage points, while in states where $185\% \leq T \leq 250\%$ it is between 10 and 22 percentage points. In states where the thresholds are above 250% the poverty line I do not find evidence that making a child eligible for Medicaid increases the chance that she receives Medicaid coverage. Higher-income families, targeted by Medicaid in more generous states, may not find beneficial to enroll their children in Medicaid because they may have better options available. This result is consistent with the quality of private health insurance being a normal good.[26]

Additionally, I take advantage of the panel structure of my dataset to perform a placebo test to check whether Medicaid participation in a period $t - \tau$ as a function of income in period t changes discontinuously at period t thresholds. If eligibility in period t is truly exogenous in the neighborhood of the threshold, then the only variable that should change discontinuously as a function of income in period t is Medicaid coverage in period t. Although there could be some correlation between income in period t and Medicaid participation in a period $t - \tau$ (because income is serially correlated), I should not observe any discontinuity in period $t - \tau$ participation at the period t threshold (i.e., eligibility in the neighborhood of the threshold in period t is exogenous and does not depend on previous Medicaid participation.) Since Medicaid participation across periods can be highly correlated, finding a discontinuity in participation in period t but not in $t - \tau$ would be a strong piece of evidence supporting the validity of the RD design (Lee and Lemieux, 2010). In Section 1.8.2.1 of Appendix B, I present graphs showing the relation between Medicaid coverage in period $t - \tau$ and family income in period t. The graphs show that participation in $t - \tau$ is negatively correlated with income in t and that it is a smooth function of family income in period t . Medicaid participation in periods $t - 2$ and $t - 3$ does not change discontinuously at the threshold although for period $t - 1$, there is a small jump at the threshold. This is likely to happen because, on the one hand, most states guarantees a minimum of six months to one year of permanence in Medicaid with independence of their family income and, on the other hand, because family income may not change substantially from one year to the other.

1.5.3 Balance of individual characteristics on either side of the thresholds

The third robustness analysis consists on checking whether children characteristics are "locally" balanced, which is an implication of the "local" randomization generated by the eligibility rule. To check for this, I run regressions of the form

$$y_{it} = \gamma_0 + \gamma_1 Eli_{it} + f_g(z_{it}) + u_{it}, \qquad t = 1997, 2002, 2007, \tag{1.12}$$

[26]These results remain the same when considering the sample that includes all years for which I can keep track family income in PSID (period 1991 and 2007). See Appendix 1.8.2.

where y_{it} are child and family characteristics not taken into account at the moment of determining Medicaid eligibility. I also consider pre-treatment variables which should not be affected by eligibility status, such as child's birth weight or mother's age at child's birth. If any of the observable characteristics changes discontinuously at the threshold, it is an indication that the eligibility rule does not generates a "local" randomization.
Table 1.3 presents the results and there are no signs of systematic discontinuous changes of characteristics at the threshold.

Table 1.3: Balance of covariates on either side of the threshold. Full sample. Years 1997, 2002, 2007.

	Bandwidth (thousands dollars)		
Dep. Var.	±30	±20	±15
Male	0.083**	0.068	0.069
	(0.038)	(0.042)	(0.045)
Black	0.016	0.025	0.018
	(0.033)	(0.036)	(0.038)
Metropolitan Area	0.044	0.066	0.061
	(0.040)	(0.042)	(0.044)
Rural Area	-0.034	-0.047	-0.037
	(0.035)	(0.038)	(0.039)
Child Birth Weight	-0.019	-0.039	-0.047
	(0.059)	(0.062)	(0.068)
Head Education (yrs)	0.047	0.095	0.057
	(0.181)	(0.189)	(0.205)
Mother age at child birth	0.481	0.436	0.523
	(0.493)	(0.527)	(0.555)
N	2163	1738	1361

Notes: Robust standard errors (in parentheses) are clustered at the family level. Each entry comes from a separate linear regression, $y_{it} = \gamma_0 + \gamma_1 Eli_{it} + f_g(z_{it}) + u_{it}$, where the dependent variable is replaced by children and family characteristics, and pre-treatment covariates. The reported coefficient is $\hat{\gamma_1}$ of equation (1.12). Each regression includes 2th order polynomial of log of income, age, and family size as well as year and state dummies.

1.6 Results

1.6.1 Contemporaneous Effects

Preventive health care utilization. Table 1.4 presents the results of the contemporaneous effects of Medicaid −equations (1.3) and (1.5)− on utilization of preventive medical care for the full sample (pooling all eligibility thresholds), and the equations (1.6) and (1.7) when allowing for heterogeneous effects for the two groups of thresholds, "low" (thresholds lower than 185%) and "high" threholds (thresholds between 185 and 2055%).[27] The table present the results for different orders of polynomials and different bandwidths. For each bandwidth, the optimal order of the polynomial is selected according to the Akaike information criterion (AIC) of model selection, which in most cases is of order 1 or 2, so I report both specifications.

[27]We showed in previous period that the discontinuity condition in the first stage is not satisfied for the group of thresholds higher than 250% of the poverty line.

The intention to treat estimates for the full sample show that making a child, with family income equal to the threshold, eligible for Medicaid increases his health care utilization by 5-10 percentage points relative to a similar child but non-eligible.[28] The IV estimates for the pooled thresholds indicate that Medicaid increases by 25-45 percentage point the probability of visiting a doctor for preventive purposes in the subpopulation of compliers −those who, made eligible for Medicaid, would enroll into the program. The results are not, however, statistically significant under all specifications.

When I allow for heterogeneous effects by threshold groups, a clear pattern emerges. The results indicate that Medicaid eligibility induces a 15-20 percentage points increase in utilization for children in the group of low thresholds, with an average effect of about 50 to 80 percentage point for the compliers.[29] However, Medicaid does not have a statistically significant impact on utilization at higher thresholds. The difference between the estimated intention to treat effects at low and high eligibility thresholds is always statistically different from zero, but not always for the LATE estimates.

Health outcomes. Tables 1.5 to 1.7 present the estimated contemporaneous effects of Medicaid on some measures of children's health: an indicator of whether the child has excellent health, obesity, and the number of school days missed due to illness. Each table presents the results for different model specifications and bandwidth choices, for the sample pooling all the thresholds, and for the subsamples of low and high thresholds. According to these results, Medicaid does not seem to have a positive effect on health in the short run for children between 5 and 18 years old. In all cases, the effects of Medicaid on these health measures are null. Even at low threshold levels the effects are null, despite Medicaid increases preventive health care utilization.

[28] Figure 1.4 in the appendix gives the graphical representation of this intention to treat effect.

[29] The estimated jump with the largest bandwidth and using a polynomial of order 1 tend to be overestimated , which gives a bias downward the LATE estimate.

Table 1.4: Contemporaneous effects of Medicaid on utilization. Children between 5 and 18 years old. *Dep. Var.: The child has visited a doctor for a routine health check-up in the last 12 months.*

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	A. Intention to treat					
One						
All Thresholds	0.080**	0.089**	0.050	0.103***	0.104***	0.057
	(0.029)	(0.033)	(0.039)	(0.033)	(0.039)	(0.043)
Low Thresholds	0.155***	0.177***	0.172***	0.184***	0.227***	0.188***
	(0.041)	(0.042)	(0.055)	(0.053)	(0.059)	(0.071)
High Thresholds	0.025	-0.015	0.005	0.044	-0.018	0.011
	(0.035)	(0.042)	(0.052)	(0.043)	(0.053)	(0.062)
Two						
All Thresholds	0.056*	0.069**	0.062	0.064	0.046	0.048
	(0.034)	(0.038)	(0.047)	(0.043)	(0.048)	(0.057)
Low Thresholds	0.152***	0.181***	0.196***	0.184**	0.203***	0.213**
	(0.050)	(0.054)	(0.070)	(0.076)	(0.078)	(0.090)
High Thresholds	-0.012	-0.016	0.029	-0.016	-0.051	-0.026
	(0.039)	(0.052)	(0.059)	(0.056)	(0.072)	(0.088)
Polynomial Order	**B. Outcome Equation. IV-RD**					
One						
All Thresholds	0.298***	0.403***	0.311	0.396***	0.454***	0.329
	(0.107)	(0.148)	(0.243)	(0.131)	(0.175)	(0.252)
Low Thresholds	0.385***	0.473***	0.611**	0.509***	0.647***	0.661**
	(0.115)	(0.144)	(0.246)	(0.159)	(0.197)	(0.284)
High Thresholds	0.169	-0.133	0.150	0.252	-0.110	0.191
	(0.178)	(0.647)	(0.583)	(0.214)	(0.445)	(0.515)
Two						
All Thresholds	0.330	0.447*	0.372	0.358	0.237	0.312
	(0.204)	(0.258)	(0.293)	(0.248)	(0.247)	(0.383)
Low Thresholds	0.490**	0.617**	0.660**	0.615**	0.673**	0.833*
	(0.201)	(0.240)	(0.297)	(0.290)	(0.311)	(0.462)
High Thresholds	-0.135	-0.108	0.338	-0.137	-0.367	-0.134
	(0.612)	(0.534)	(0.571)	(0.541)	(0.633)	(0.612)
Different Polynomial at each side	N	N	N	Y	Y	Y
N	2613	1904	1486	2613	1904	1486

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The intention to treat estimates in each column, when not allowing for heterogeneous effects ("all thresholds" pooled), come from the following model: $y_{it} = \alpha + \theta Eli_{it} + f_g(z_{it}) + f_g(z_{it}) \times Eli_{it} + u_{it}$. The IV-RD estimates for the full sample ("all thresholds") in each column come from the following model: $y_{it} = \alpha + \beta M_{it} + k_g(z_{it}) + k_g(z_{it}) \times Eli_{it} + u_{it}$, where eligibility instruments for Medicaid coverage. The intention to treat estimates in each column, when allowing for heterogeneous effects by threshold groups ("low" and "high" thresholds), come from the following model: $y_{it} = \alpha_0 + \theta_0 Eli_{0it} + \theta_1 Eli_{1it} + \alpha_1 T_{1it} + f_g(z_{it}) + f_g(z_{it}) \times T_{1it} + f_g(z_{it}) \times Eli_{1it} + f_g(z_{it}) \times T_{1it} \times Eli_{1it} + u_{it}$, where Eli_{0it} indicates eligibility in a low threshold state, and Eli_{1it} indicates eligibility in a high threshold state. The IV-RD estimates in each column come from the following model: $y_{it} = \alpha_0 + \beta_0 M_{0it} + \beta_1 M_{1it} + \alpha_1 T_{1it} + k_g(z_{it}) + k_g(z_{it}) \times T_{1it} + k_g(z_{it}) \times Eli_{1it} + k_g(z_{it}) \times Eli_{1it} \times T_{1it} + u_{it}$, where eligibility instruments for Medicaid coverage. The first three columns do not allow for different polynomials at each side of the cutoff.

Table 1.5: Contemporaneous effects of Medicaid on the probability of being in Excellent Health. Children between 5 and 18 years old.

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	A. Intention to treat					
One						
All Thresholds	-0.053	-0.012	-0.011	-0.062*	-0.018	-0.015
	(0.034)	(0.036)	(0.045)	(0.038)	(0.042)	(0.050)
Low Thresholds	-0.071	-0.023	-0.001	-0.100	-0.031	-0.010
	(0.053)	(0.053)	(0.069)	(0.062)	(0.065)	(0.076)
High Thresholds	-0.041	-0.018	-0.054	-0.016	-0.023	-0.050
	(0.039)	(0.053)	(0.065)	(0.051)	(0.065)	(0.077)
Two						
All Thresholds	-0.036	0.003	-0.029	-0.009	-0.006	-0.064
	(0.038)	(0.042)	(0.055)	(0.049)	(0.058)	(0.071)
Low Thresholds	-0.046	0.009	-0.010	-0.022	0.031	-0.014
	(0.064)	(0.066)	(0.087)	(0.081)	(0.088)	(0.115)
High Thresholds	-0.040	-0.044	-0.066	-0.060	-0.118	-0.203**
	(0.048)	(0.064)	(0.072)	(0.071)	(0.091)	(0.102)
Polynomial Order	**B. Outcome Equation. IV-RD**					
One						
All Thresholds	-0.204	-0.062	-0.071	-0.243*	-0.094	-0.092
	(0.126)	(0.182)	(0.301)	(0.145)	(0.214)	(0.302)
Low Thresholds	-0.232	-0.098	-0.039	-0.364	-0.148	-0.069
	(0.164)	(0.201)	(0.398)	(0.228)	(0.302)	(0.459)
High Thresholds	-0.176	-0.129	-0.421	-0.059	-0.124	-0.328
	(0.171)	(0.454)	(0.591)	(0.201)	(0.373)	(0.518)
Two						
All Thresholds	-0.217	0.018	-0.194	-0.049	-0.033	-0.391
	(0.237)	(0.282)	(0.374)	(0.273)	(0.298)	(0.443)
Low Thresholds	-0.263	0.021	-0.063	-0.151	0.120	0.005
	(0.353)	(0.337)	(0.627)	(0.458)	(0.466)	(0.702)
High Thresholds	-0.312	-0.300	-0.407	-0.404	-0.913	-1.445
	(0.441)	(0.482)	(0.521)	(0.498)	(0.809)	(1.125)
Different Polynomial at each side	N	N	N	Y	Y	Y
N	2178	1581	1242	2178	1581	1242

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The intention to treat estimates in each column, when not allowing for heterogeneous effects ("all thresholds" pooled), come from the following model: $y_{it} = \alpha + \theta Eli_{it} + f_g(z_{it}) + f_g(z_{it}) \times Eli_{it} + u_{it}$. The IV-RD estimates for the full sample ("all thresholds") in each column come from the following model: $y_{it} = \alpha + \beta M_{it} + k_g(z_{it}) + k_g(z_{it}) \times Eli_{it} + u_{it}$, where eligibility instruments for Medicaid coverage. The intention to treat estimates in each column, when allowing for heterogeneous effects by threshold groups ("low" and "high" thresholds), come from the following model: $y_{it} = \alpha_0 + \theta_0 Eli_{0it} + \theta_1 Eli_{1it} + \alpha_1 T_{1it} + f_g(z_{it}) + f_g(z_{it}) \times T_{1it} + f_g(z_{it}) \times Eli_{1it} + f_g(z_{it}) \times T_{1it} \times Eli_{1it} + u_{it}$, where Eli_{0it} indicates eligibility in a low threshold state, and Eli_{1it} indicates eligibility in a high threshold state. The IV-RD estimates in each column come from the following model: $y_{it} = \alpha_0 + \beta_0 M_{0it} + \beta_1 M_{1it} + \alpha_1 T_{1it} + k_g(z_{it}) + k_g(z_{it}) \times T_{1it} + k_g(z_{it}) \times Eli_{1it} + k_g(z_{it}) \times Eli_{1it} \times T_{1it} + u_{it}$, where eligibility instruments for Medicaid coverage. The first three columns do not allow for different polynomials at each side of the cutoff.

Table 1.6: Contemporaneous effects of Medicaid on Obesity. Children between 5 and 18 years old.

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	A. Intention to treat					
One						
All Thresholds	0.003	-0.059	-0.011	-0.020	-0.059	-0.016
	(0.031)	(0.040)	(0.043)	(0.036)	(0.043)	(0.049)
Low Thresholds	0.005	-0.058	0.008	-0.015	-0.055	0.009
	(0.050)	(0.062)	(0.064)	(0.056)	(0.068)	(0.074)
High Thresholds	0.004	-0.078	-0.043	-0.037	-0.088	-0.053
	(0.039)	(0.056)	(0.068)	(0.046)	(0.057)	(0.071)
Two						
All Thresholds	-0.039	-0.055	-0.028	-0.058	-0.037	-0.045
	(0.037)	(0.045)	(0.053)	(0.049)	(0.058)	(0.067)
Low Thresholds	-0.055	-0.052	-0.011	-0.017	0.008	-0.022
	(0.059)	(0.069)	(0.081)	(0.083)	(0.094)	(0.112)
High Thresholds	-0.031	-0.061	-0.046	-0.094	-0.020	-0.015
	(0.048)	(0.062)	(0.075)	(0.065)	(0.086)	(0.102)
Polynomial Order	**B. Outcome Equation. IV-RD**					
One						
All Thresholds	0.012	-0.423	-0.099	-0.606***	-0.625	-0.864
	(0.117)	(0.323)	(0.400)	(0.209)	(0.407)	(0.783)
Low Thresholds	-0.015	-0.055	0.009	-0.062	-0.276	0.093
	(0.056)	(0.068)	(0.074)	(0.221)	(0.401)	(0.677)
High Thresholds	-0.037	-0.088	-0.053	-0.193	-1.115	-1.277
	(0.046)	(0.057)	(0.071)	(0.245)	(1.122)	(2.725)
Two						
All Thresholds	-0.332	-0.595	-0.377	-0.880	-1.479	-2.284
	(0.333)	(0.561)	(0.729)	(0.866)	(1.611)	(2.673)
Low Thresholds	-0.017	0.008	-0.022	-0.360	-0.048	0.052
	(0.083)	(0.094)	(0.112)	(0.852)	(1.123)	(4.703)
High Thresholds	-0.094	-0.020	-0.015	-1.657	-1.304	1.993
	(0.065)	(0.086)	(0.102)	(2.327)	(8.197)	(43.402)
Different Polynomial at each side	N	N	N	Y	Y	Y
N	2178	1581	1242	2178	1581	1242

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The intention to treat estimates in each column, when not allowing for heterogeneous effects ("all thresholds" pooled), come from the following model: $y_{it} = \alpha + \theta Eli_{it} + f_g(z_{it}) + f_g(z_{it}) \times Eli_{it} + u_{it}$. The IV-RD estimates for the full sample ("all thresholds") in each column come from the following model: $y_{it} = \alpha + \beta M_{it} + k_g(z_{it}) + k_g(z_{it}) \times Eli_{it} + u_{it}$, where eligibility instruments for Medicaid coverage. The intention to treat estimates in each column, when allowing for heterogeneous effects by threshold groups ("low" and "high" thresholds), come from the following model: $y_{it} = \alpha_0 + \theta_0 Eli_{0it} + \theta_1 Eli_{1it} + \alpha_1 T_{1it} + f_g(z_{it}) + f_g(z_{it}) \times T_{1it} + f_g(z_{it}) \times Eli_{1it} + f_g(z_{it}) \times T_{1it} \times Eli_{1it} + u_{it}$, where Eli_{0it} indicates eligibility in a low threshold state, and Eli_{1it} indicates eligibility in a high threshold state. The IV-RD estimates in each column come from the following model: $y_{it} = \alpha_0 + \beta_0 M_{0it} + \beta_1 M_{1it} + \alpha_1 T_{1it} + k_g(z_{it}) + k_g(z_{it}) \times T_{1it} + k_g(z_{it}) \times Eli_{1it} + k_g(z_{it}) \times Eli_{1it} \times T_{1it} + u_{it}$, where eligibility instruments for Medicaid coverage. The first three columns do not allow for different polynomials at each side of the cutoff.

Table 1.7: Contemporaneous effects of Medicaid on Number of School Days Missed due to Illness . Children between 5 and 18 years old.

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	A. Intention to treat					
One						
All Thresholds	-0.097	-0.198	-1.107*	-0.133	-0.279	-0.715
	(0.333)	(0.443)	(0.596)	(0.349)	(0.398)	(0.484)
Low Thresholds	0.550	0.635	-1.198	0.545	0.378	-0.981
	(0.671)	(0.749)	(0.803)	(0.725)	(0.793)	(0.795)
High Thresholds	-0.131	-0.280	0.142	0.202	0.004	0.714
	(0.361)	(0.458)	(0.639)	(0.457)	(0.524)	(0.828)
Two						
All Thresholds	-0.753**	-0.778*	-0.267	-0.133	-0.279	-0.715
	(0.380)	(0.440)	(0.496)	(0.349)	(0.398)	(0.484)
Low Thresholds	-0.713	-0.533	-0.469	-1.102	-1.130	-1.016
	(0.665)	(0.622)	(0.799)	(0.887)	(0.911)	(1.207)
High Thresholds	-0.496	-0.087	0.951	0.416	0.951	1.232
	(0.393)	(0.519)	(0.869)	(0.603)	(0.782)	(0.862)
Polynomial Order	**B. Outcome Equation. IV-RD**					
One						
All Thresholds	-0.373	-0.997	-7.400	-1.789	-2.333	1.688
	(1.291)	(2.306)	(4.453)	(2.530)	(2.661)	(4.225)
Low Thresholds	1.649	2.183	-7.018	2.008	1.751	-5.765
	(2.040)	(2.810)	(5.766)	(2.625)	(3.696)	(5.633)
High Thresholds	-0.455	-2.298	1.603	0.770	-0.072	5.097
	(1.736)	(4.033)	(5.734)	(1.793)	(3.033)	(6.123)
Two						
All Thresholds	-0.332	-0.595	-0.377	-0.880	-1.479	-2.284
	(0.333)	(0.561)	(0.729)	(0.866)	(1.611)	(2.673)
Low Thresholds	-3.997	-2.797	-3.513	-5.882	-5.440	-6.528
	(3.822)	(3.474)	(6.645)	(5.777)	(5.467)	(8.902)
High Thresholds	-3.836	-0.444	6.125	3.559	8.208	9.814
	(3.793)	(3.884)	(6.443)	(4.558)	(8.091)	(9.876)
Different Polynomial at each side	N	N	N	Y	Y	Y
N	2178	1581	1242	2178	1581	1242

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The intention to treat estimates in each column, when not allowing for heterogeneous effects ("all thresholds" pooled), come from the following model: $y_{it} = \alpha + \theta Eli_{it} + f_g(z_{it}) + f_g(z_{it}) \times Eli_{it} + u_{it}$. The IV-RD estimates for the full sample ("all thresholds") in each column come from the following model: $y_{it} = \alpha + \beta M_{it} + k_g(z_{it}) + k_g(z_{it}) \times Eli_{it} + u_{it}$, where eligibility instruments for Medicaid coverage. The intention to treat estimates in each column, when allowing for heterogeneous effects by threshold groups ("low" and "high" thresholds), come from the following model: $y_{it} = \alpha_0 + \theta_0 Eli_{0it} + \theta_1 Eli_{1it} + \alpha_1 T_{1it} + f_g(z_{it}) + f_g(z_{it}) \times T_{1it} + f_g(z_{it}) \times Eli_{1it} + f_g(z_{it}) \times T_{1it} \times Eli_{1it} + u_{it}$, where Eli_{0it} indicates eligibility in a low threshold state, and Eli_{1it} indicates eligibility in a high threshold state. The IV-RD estimates in each column come from the following model: $y_{it} = \alpha_0 + \beta_0 M_{0it} + \beta_1 M_{1it} + \alpha_1 T_{1it} + k_g(z_{it}) + k_g(z_{it}) \times T_{1it} + k_g(z_{it}) \times Eli_{1it} + k_g(z_{it}) \times Eli_{1it} \times T_{1it} + u_{it}$, where eligibility instruments for Medicaid coverage. The first three columns do not allow for different polynomials at each side of the cutoff.

1.6.2 Lagged Effects on Health

1.6.2.1 Full Sample

Now I turn to the analysis of the cumulative effects of Medicaid in the medium run, looking first at the effects on the sample pooling all thresholds. Tables 1.8 and 1.9 report the cumulative IIT estimates, which capture the effect of making a child randomly eligible for Medicaid in a given period on the probability of being in excellent health and obesity after τ periods −equation (1.8).[30] These ITT estimates identify the effects of eligibility in one moment of time on future outcomes, without controlling for behavioral changes between the period of eligibility and the period in which outcomes are measured. Thus, IIT estimates reflect accumulated effects as shown in equation (1.9).[31]

Results indicate that Medicaid has a negative effect on the probability of being in excellent health after one and two years (and the effect then vanishes) and a positive effect on the probability of being obese after two years (and the effect then vanishes). Making a child eligible for Medicaid decreases the likelihood of being in excellent health in about 7-11 percentage points after one year, relative to a similar non eligible child. This negative effect still persists after two years and it is almost of the same magnitude. Similarly, making a child eligible for Medicaid increases the probability of being obese in 10-14 percentage points after two years, relative to a similar child not eligible for Medicaid.[32]

1.6.2.2 Threshold groups

The finding that Medicaid has a differential impact on preventive health care utilization by threshold groups, rises the question whether it also has a differential impact on health outcomes in the medium run. Allowing for heterogeneous effects of Medicaid on the probability of being in excellent health by threshold groups, we can observe from Table (1.10) that the negative effects after one and two years are, however, present at both low and high threshold groups, although the results are weaker (the effects are not statistically significant in many specifications but the magnitude of the effects are similar across models.) The negative effects tends two vanish afterwards, since the effects are not significant for neither group after 3 years onwards.

Similarly, results in Table 1.11 show that the same pattern emerges for the effect of Medicaid on the probability of being obese after two years. Medicaid has a positive impact on (it increases) the probability of being obese at both low and high thresholds after two years, and the effects vanish afterwards.

[30]The lagged effects on school days missed due to illness are present in table 1.15 in the Appendix. No effects are found on this variable.

[31]I do not report the IV estimates because they tend to be more imprecisely estimated. However, ITT effects are lower bounds for the average treatment effects.

[32]I also find, consistent with this result, that Medicaid increases child BMI after two years.

Table 1.8: Lagged cumulative intention to treat effects of Medicaid on Excellent Health. Children between 5 and 18 years old. Full sample (all thresholds pooled).

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	Eligible One Year Before (θ_1)					
One	-0.096***	-0.108***	-0.073*	-0.093***	-0.115***	-0.090***
	(0.031)	(0.036)	(0.042)	(0.035)	(0.042)	(0.047)
Two	-0.076**	-0.082**	-0.079	-0.056	-0.071	-0.065
	(0.036)	(0.040)	(0.051)	(0.048)	(0.053)	(0.063)
Polynomial Order	**Eligible Two Years Before (θ_2)**					
One	-0.070*	-0.081**	-0.081*	-0.067*	-0.078*	-0.089*
	(0.036)	(0.040)	(0.047)	(0.040)	(0.044)	(0.051)
Two	-0.067	-0.090*	-0.093*	-0.054	-0.065	-0.040
	(0.043)	(0.048)	(0.054)	(0.053)	(0.059)	(0.066)
	Eligible Three Years Before (θ_3)					
One	-0.048	-0.062	-0.060	-0.032	-0.041	-0.063
	(0.033)	(0.038)	(0.045)	(0.040)	(0.044)	(0.051)
Two	-0.048	-0.075	-0.071	-0.030	-0.053	-0.022
	(0.039)	(0.046)	(0.052)	(0.047)	(0.051)	(0.062)
	Eligible Four Years Before (θ_4)					
One	-0.052	-0.048	0.001	-0.022	-0.027	0.012
	(0.034)	(0.037)	(0.043)	(0.040)	(0.045)	(0.051)
Two	0.012	0.006	-0.010	0.023	-0.013	-0.005
	(0.041)	(0.043)	(0.054)	(0.053)	(0.060)	(0.069)
	Eligible Five Years Before (θ_5)					
One	-0.065*	-0.022	-0.042	-0.045	0.023	-0.008
	(0.034)	(0.036)	(0.044)	(0.038)	(0.042)	(0.050)
Two	-0.031	-0.008	-0.030	0.020	0.022	0.021
	(0.039)	(0.043)	(0.052)	(0.049)	(0.055)	(0.064)
Different Polynomial at each side	N	N	N	Y	Y	Y

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The lagged intention to treat estimates in each column come from the following model: $y_{it} = \alpha + \theta_\tau Eli_{i,t-\tau} + f_g(z_{i,t-\tau}) + f_g(z_{i,t-\tau}) \times Eli_{i,t-\tau} + u_{it}$. The first three columns do not allow for different polynomials at each side of the cutoff.

Table 1.9: Lagged cumulative intention to treat effects of Medicaid on Obesity. Children between 5 and 18 years old. Full sample (all thresholds pooled).

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	Eligible One Year Before (θ_1)					
One	0.024	0.038	0.039	0.030	0.033	0.027
	(0.025)	(0.027)	(0.035)	(0.028)	(0.032)	(0.038)
Two	0.031	0.043	0.042	0.019	0.020	0.066
	(0.029)	(0.033)	(0.039)	(0.037)	(0.042)	(0.048)
Polynomial Order	**Eligible Two Years Before (θ_2)**					
One	0.095***	0.108***	0.130***	0.098***	0.105***	0.129***
	(0.031)	(0.033)	(0.038)	(0.032)	(0.036)	(0.042)
Two	0.119***	0.125***	0.142***	0.136***	0.130***	0.081
	(0.034)	(0.038)	(0.044)	(0.042)	(0.047)	(0.055)
	Eligible Three Years Before (θ_3)					
One	0.053	0.042	0.043	0.053	0.048	0.021
	(0.032)	(0.035)	(0.042)	(0.035)	(0.041)	(0.047)
Two	0.049	0.041	0.050	-0.017	-0.039	-0.033
	(0.037)	(0.042)	(0.050)	(0.049)	(0.054)	(0.063)
	Eligible Four Years Before (θ_4)					
One	0.010	0.002	0.017	-0.022	-0.027	0.012
	(0.025)	(0.027)	(0.033)	(0.040)	(0.045)	(0.051)
Two	0.006	0.006	-0.001	0.023	-0.013	-0.005
	(0.032)	(0.034)	(0.043)	(0.053)	(0.060)	(0.069)
	Eligible Five Years Before (θ_5)					
One	0.027	0.019	0.024	0.038	0.044	0.072**
	(0.024)	(0.027)	(0.034)	(0.028)	(0.031)	(0.037)
Two	0.025	0.020	0.076	0.060	0.079	0.067
	(0.030)	(0.031)	(0.037)	(0.037)	(0.041)	(0.050)
Different Polynomial at each side	N	N	N	Y	Y	Y

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The lagged intention to treat estimates in each column come from the following model: $y_{it} = \alpha + \theta_\tau Eli_{i,t-\tau} + f_g(z_{i,t-\tau}) + f_g(z_{i,t-\tau}) \times Eli_{i,t-\tau} + u_{it}$. The first three columns do not allow for different polynomials at each side of the cutoff.

Table 1.10: Lagged cumulative intention to treat effect of Medicaid on Excellent Health. Children between 5 and 18 years old.

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	Eligible One Year Before (θ_1)					
One						
Low Thresholds	-0.102**	-0.094**	-0.070	-0.085	-0.122**	-0.095
	(0.047)	(0.049)	(0.060)	(0.056)	(0.064)	(0.073)
High Thresholds	-0.087**	-0.089**	-0.089	-0.082*	-0.067	-0.088
	(0.036)	(0.049)	(0.059)	(0.047)	(0.058)	(0.066)
Two						
Low Thresholds	-0.053	-0.062	-0.068	-0.031	-0.075	-0.028
	(0.053)	(0.059)	(0.075)	(0.079)	(0.087)	(0.108)
High Thresholds	-0.093**	-0.051	-0.067	-0.063	-0.075	-0.089
	(0.044)	(0.057)	(0.065)	(0.063)	(0.078)	(0.096)
Polynomial Order	**Eligible Two Years Before (θ_2)**					
One						
Low Thresholds	-0.045	-0.070	-0.096	-0.076	-0.094	-0.129*
	(0.055)	(0.058)	(0.066)	(0.053)	(0.059)	(0.067)
High Thresholds	-0.080**	-0.060	-0.024	-0.060	-0.008	0.013
	(0.042)	(0.054)	(0.065)	(0.052)	(0.060)	(0.069)
Two						
Low Thresholds	-0.061	-0.119*	-0.153*	-0.108	-0.150*	-0.155*
	(0.063)	(0.070)	(0.078)	(0.073)	(0.083)	(0.094)
High Thresholds	-0.047	-0.004	0.014	0.025	0.057	0.052
	(0.052)	(0.063)	(0.072)	(0.068)	(0.083)	(0.099)
Polynomial Order	**Eligible Three Years Before (θ_3)**					
One						
Low Thresholds	-0.037	-0.058	-0.096	-0.014	-0.051	-0.101
	(0.046)	(0.052)	(0.060)	(0.058)	(0.067)	(0.073)
High Thresholds	-0.012	0.049	0.043	0.044	0.064	0.081
	(0.054)	(0.068)	(0.083)	(0.071)	(0.087)	(0.096)
Two						
Low Thresholds	-0.045	-0.089	-0.062	-0.075	-0.180	-0.085
	(0.054)	(0.058)	(0.074)	(0.079)	(0.086)	(0.099)
High Thresholds	0.056	0.040	0.081	0.077	0.070	0.135
	(0.068)	(0.083)	(0.094)	(0.102)	(0.125)	(0.146)
Polynomial Order	**Eligible Four Years Before (θ_4)**					
One						
Low Thresholds	-0.063*	-0.053	0.005	-0.019	-0.038	-0.003
	(0.038)	(0.041)	(0.050)	(0.052)	(0.059)	(0.066)
High Thresholds	-0.023	0.024	0.043	0.020	0.060	0.109
	(0.051)	(0.063)	(0.079)	(0.064)	(0.076)	(0.090)
Two						
Low Thresholds	0.020	0.007	-0.028	-0.010	-0.068	-0.071
	(0.048)	(0.051)	(0.065)	(0.071)	(0.079)	(0.091)
High Thresholds	0.037	0.066	0.098	0.146*	0.099	0.004
	(0.063)	(0.079)	(0.091)	(0.088)	(0.106)	(0.125)
Polynomial Order	**Eligible Five Years Before (θ_5)**					
One						
Low Thresholds	-0.063*	-0.032	-0.070	-0.052	0.006	-0.056
	(0.035)	(0.037)	(0.047)	(0.045)	(0.049)	(0.059)
High Thresholds	-0.075	0.065	0.120	0.008	0.150**	0.204**
	(0.050)	(0.068)	(0.090)	(0.067)	(0.081)	(0.097)
Two						
Low Thresholds	-0.034	-0.026	-0.081	-0.019	-0.029	-0.056
	(0.041)	(0.044)	(0.055)	(0.059)	(0.065)	(0.078)
High Thresholds	-0.022	0.141	0.208	0.199	0.147	0.048
	(0.061)	(0.088)	(0.098)	(0.088)	(0.114)	(0.134)
Different Polynomial at each side	N	N	N	Y	Y	Y

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The lagged intention to treat estimates in each column come from the following model: $y_{it} = \alpha_0 + \theta_{low,\tau} Eli_{0i,t-\tau} + \theta_{high,\tau} Eli_{1i,t-\tau} + \alpha_1 T_{1i,t-\tau} + f_g(z_{i,t-\tau}) + f_g(z_{i,t-\tau}) \times T_{1i,t-\tau} + f_g(z_{i,t-\tau}) \times Eli_{1i,t-\tau} + f_g(z_{i,t-\tau}) \times T_{1i,t-\tau} \times Eli_{1i,t-\tau} + u_{it}$, where Eli_{0it} indicates eligibility in a low threshold state, and Eli_{1it} indicates eligibility in a high threshold state. The first three columns do not allow for different polynomials at each side of the cutoff.

Table 1.11: Lagged cumulative intention to treat effect of Medicaid on Obesity. Children between 5 and 18 years old.

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	Eligible One Year Before (θ_1)					
One						
Low Thresholds	0.021	0.049	0.062	0.038	0.063	0.054
	(0.039)	(0.039)	(0.054)	(0.045)	(0.051)	(0.061)
High Thresholds	0.022	0.021	0.027	0.010	-0.004	0.014
	(0.031)	(0.042)	(0.050)	(0.037)	(0.045)	(0.051)
Two						
Low Thresholds	0.043	0.068	0.082	0.055	0.071	0.136
	(0.046)	(0.051)	(0.063)	(0.063)	(0.069)	(0.078)
High Thresholds	0.014	0.004	0.007	-0.003	0.012	0.034
	(0.037)	(0.048)	(0.055)	(0.048)	(0.060)	(0.076)
Polynomial Order	**Eligible Two Years Before (θ_2)**					
One						
Low Thresholds	0.097**	0.105**	0.115**	0.085*	0.113**	0.134**
	(0.049)	(0.052)	(0.058)	(0.045)	(0.049)	(0.057)
High Thresholds	0.092**	0.109**	0.118**	0.087**	0.084*	0.070
	(0.038)	(0.045)	(0.056)	(0.042)	(0.047)	(0.053)
Two						
Low Thresholds	0.110**	0.138**	0.165**	0.164***	0.165**	0.119
	(0.054)	(0.060)	(0.069)	(0.059)	(0.068)	(0.079)
High Thresholds	0.122***	0.103*	0.069	0.061	0.032	-0.017
	(0.044)	(0.053)	(0.061)	(0.053)	(0.065)	(0.081)
Polynomial Order	**Eligible Three Years Before (θ_3)**					
One						
Low Thresholds	0.022	0.013	-0.006	0.021	0.018	-0.022
	(0.041)	(0.044)	(0.054)	(0.051)	(0.058)	(0.068)
High Thresholds	0.091*	0.107**	0.094	0.069	0.101	0.048
	(0.048)	(0.060)	(0.069)	(0.054)	(0.067)	(0.072)
Two						
Low Thresholds	0.009	0.000	0.008	-0.077	-0.114	-0.084
	(0.048)	(0.052)	(0.067)	(0.070)	(0.078)	(0.092)
High Thresholds	0.079	0.106	0.015	0.045	-0.009	-0.051
	(0.059)	(0.070)	(0.075)	(0.078)	(0.096)	(0.110)
Polynomial Order	**Eligible Four Years Before (θ_4)**					
One						
Low Thresholds	0.016	0.017	0.037	-0.019	-0.019	0.017
	(0.032)	(0.033)	(0.043)	(0.042)	(0.049)	(0.058)
High Thresholds	-0.011	-0.018	-0.037	-0.018	-0.014	-0.020
	(0.043)	(0.051)	(0.066)	(0.052)	(0.059)	(0.069)
Two						
Low Thresholds	0.023	0.029	0.020	-0.007	-0.012	-0.052
	(0.042)	(0.044)	(0.057)	(0.062)	(0.070)	(0.083)
High Thresholds	-0.031	-0.036	-0.020	-0.052	-0.057	-0.039
	(0.053)	(0.064)	(0.073)	(0.073)	(0.090)	(0.108)
Polynomial Order	**Eligible Five Years Before (θ_5)**					
One						
Low Thresholds	0.021	0.018	0.022	0.026	0.045	0.075*
	(0.026)	(0.030)	(0.038)	(0.033)	(0.037)	(0.043)
High Thresholds	0.059	0.059	0.041	0.100*	0.058	0.056
	(0.040)	(0.055)	(0.074)	(0.054)	(0.063)	(0.076)
Two						
Low Thresholds	0.014	0.020	0.083	0.041	0.086*	0.075
	(0.034)	(0.036)	(0.044)	(0.045)	(0.049)	(0.059)
High Thresholds	0.072	0.042	0.055	0.080	0.090	0.057
	(0.050)	(0.070)	(0.079)	(0.075)	(0.094)	(0.110)
Different Polynomial at each side	N	N	N	Y	Y	Y

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The lagged intention to treat estimates in each column come from the following model: $y_{it} = \alpha_0 + \theta_{low,\tau} Eli_{0i,t-\tau} + \theta_{high,\tau} Eli_{1i,t-\tau} + \alpha_1 T_{1i,t-\tau} + f_g(z_{i,t-\tau}) + f_g(z_{i,t-\tau}) \times T_{1i,t-\tau} + f_g(z_{i,t-\tau}) \times Eli_{1i,t-\tau} + f_g(z_{i,t-\tau}) \times T_{1i,t-\tau} \times Eli_{1i,t-\tau} + u_{it}$, where Eli_{0it} indicates eligibility in a low threshold state, and Eli_{1it} indicates eligibility in a high threshold state. The first three columns do not allow for different polynomials at each side of the cutoff.

1.6.3 Channels

While the findings so far indicate that Medicaid increases contemporaneously the utilization of preventive health care services, the effects over health are null in the short run or even negative in the medium run. Lagged negative effects are observed after one and two years, and them vanish. A possible explanation for this somehow puzzling results could be that these negative effects are just due to a "perception" effect and they do not reflect a real change in children's health. For example, since the excellent health indicator is a subjective measure reported by children's caregivers it can be argued that given that Medicaid induces more contacts with physicians, parents become more aware of certain health problems their children already had, and this is the effect which is captured. This explanation may be only plausible to explain the negative effects of Medicaid on the probability of being in excellent health for children in the low thresholds group. However, it is not so clear that this mechanism can explain the negative effects on the high threshold group (for which Medicaid does not induce higher use of preventive medical services) neither the negative effects on obesity, which is a more objective measure of children's health.

A second channel consistent with this result is the "quality" channel, that is, differences in the quality of health care (real or perceived by parents) the child has access to through Medicaid relative to the counterfactual situation without Medicaid. If Medicaid induces targeted families to drop a private health insurance (crowding out effect), and if Medicaid's quality is lower than their previous private option, then this may translate to a negative impact on their children's health. Although PSID and CDS datasets do not provide information about the quality of private insurance to directly test whether the quality channel is operating, they do have information in some periods about private insurance coverage. This allows me to check the presence of a crowding out effect.

Table 1.12 presents the contemporaneous effects of Medicaid on the probability of having a private insurance. The results for the full sample suggest that making a child eligible for Medicaid reduces the probability of having private coverage by 12-20 percentage points relative to a similar non-eligible child.[33] The crowding out effect seems to be larger at lower thresholds (between 18-29 percentage points decrease), although in most cases they are not statistically different to that of higher thresholds. Estimates at high thresholds are quite imprecise, but the magnitude of the effect remains almost the same across specifications, ranging from 3 to 9 percentage points reduction in the probability of having private coverage. Provided that a crowding out effect is present, it could be possible then that quality differentials lead to the negative effects of Medicaid on children's health outcomes. There is indirect evidence consistent with the quality differential between Medicaid and private

[33]Excluding the results for the specification with polynomial of order 1 in the largest sample (bandwidth +-30).

Table 1.12: Crowding out effect at the threshold. Years 1997, 2002, 2007.

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
One						
All Thresholds	-0.257***	-0.201***	-0.116**	-0.220***	-0.165***	-0.126*
	(0.049)	(0.050)	(0.059)	(0.047)	(0.049)	(0.064)
Low Thresholds	-0.366***	-0.293***	-0.224***	-0.286***	-0.220***	-0.199**
	(0.072)	(0.064)	(0.076)	(0.068)	(0.071)	(0.086)
High Thresholds	-0.181***	-0.078	-0.079	-0.172***	-0.094	-0.088
	(0.050)	(0.060)	(0.076)	(0.058)	(0.069)	(0.082)
Two						
All Thresholds	-0.133***	-0.129***	-0.106	-0.116**	-0.130**	-0.041
	(0.043)	(0.047)	(0.068)	(0.054)	(0.062)	(0.086)
Low Thresholds	-0.195***	-0.199***	-0.175*	-0.286***	-0.220***	-0.199**
	(0.068)	(0.073)	(0.097)	(0.068)	(0.071)	(0.086)
High Thresholds	-0.092*	-0.080	-0.093	-0.053	-0.058	-0.031
	(0.055)	(0.074)	(0.086)	(0.076)	(0.095)	(0.112)
Different Polynomial at each side	N	N	N	Y	Y	Y
N	2163	1738	1361	2163	1738	1361

insurances hypothesis. For example, according to the annual State of Health Care Quality Report of the National Committee for Quality Assurance (NCQA), Medicaid plans tend to perform worse, on average, than commercial plans in some important quality dimensions, such as whether physicians regularly keep track of children's health by calculating their BMI, or whether during the visit the physicians give counseling about nutritions issues and recommended levels of physical activity to maintain children's health.[34] During 2010, among children between 2 to 17 years old enrolled in Medicaid and who had at least one outpatient visit with a primary care physician during the year, 32.5% received counseling about recommended physical activity levels to mantain health, while this percentage was 36.5% for children in private plans. Also, physicians documented the child's BMI in 30.3% of the Medicaid enrollees visits, while they documented it in 35.4% of the privately insured chidren's visits. There were no differenreces in the percentage of cases they give counseling about nutritions issues to Medicaid and privately insured partients (NCQA, 2010).

Other measure of quality is whether physicians follow the recommended protocols to treat certain illnesses such as pharyngitis or asthma.[35] According to the NCQA report, the percentage of children between 2 and 18 years old who were diagnosed with pharyngitis and received an appropriate testing was 59.0% in Medicaid versus 74.7% in commercial plans. Furthermore, the percentage of Medicaid patients with persistent asthma who were prescribed medications acceptable as primary therapy for long-term control of asthma was lower than that for patients enrolled in commercial plans (89.6% in Medicaid versus 96.4%

[34]The State of Health Care Quality report is produced annually by NCQA to monitor and report on performance trends over time, track variations in patterns of care and provide recommendations for future quality improvement. This report shows indicators coming from The Healthcare Effectiveness Data and Information Set (HEDIS), a tool used by more than 90 percent of America's health plans to measure performance on important dimensions of care and service.

[35]The recommended testing for pharyngitis consist on giving an antibiotic and performing a Group A streptococcus test for the episode.

in commercial plans, for children between 5 and 9 years old; and 87.0% in Medicaid versus 92.9% in commercial plans, for children between 10 and 17 years old) (NCQA, 2007).

There is some research also providing evidence of a lower quality of Medicaid relative to private insurances in other dimensions. For instance, the amount of time that a doctor spends on average with a Medicaid patient during a visit is lower than that for a privately insured patient, as shown by Decker (2007). She finds that in states where Medicaid pays lower fees the amount of time a doctor spends with Medicaid patients is lower relative to privately insured patients. Also, physicians are less likely to want to see a Medicaid patient. Hence, a Medicaid beneficiary not only finds more difficult to locate a physician willing to see him, but also the quality of care he receives, measured by the duration of the visit, is also lower than that received by a privately insured patient. Cunningham and O'Malley (2009) also find that Medicaid reimbursement delay affects physicians' willingness to accept Medicaid patients. They show that delays in reimbursement can offset the effects of high Medicaid fees, thereby lowering participation to levels that are closer to those in states with relatively low fees.

There are some limitations to the extent to which the quality channel is the one operating. On the one hand, the evidence available about the differences in quality applies to the average Medicaid population, but not to the marginal groups affected at the eligibility thresholds. The fact that I estimate the effects across different threshold levels (i.e, across different family income levels) gives more generality to my results, but still they do not apply to the whole Medicaid population. On the other hand, the quality channel is more likely to be applicable only to children in higher income families. The reasons are twofold. First, higher income families are more likely to have private insurance coverage, as shown in the Table 3.2 of Section 3.3. Indeed, the data show it is more likely that a non-eligible child has private coverage the higher the family income. Second, the quality of care families have access to through a private insurance may increase with income, i.e., health insurance quality is a normal good. Then, it is more likely that a non-eligible child has a better quality private coverage the higher the income. This intuition is consistent with the finding in Section 1.5.2 that families at high thresholds are slightly less likely than families at low thresholds to enroll their children in Medicaid despite being eligible. However, if a higher income family is induced to drop a private insurance in favor of Medicaid, this may imply a drop in health care quality and may have a negative impact on their child's health.

There are at least two other channel through which the crowding out effect induced by Medicaid may generate a negative impact on children's health: waiting times and consumption effects effects. First, it may take some time until an eligible child have effective access to Medicaid coverage and, hence, he may experience a transitory period of being uninsured which entails a higher health vulnerability. Since 1997, when the Medicaid program for children was augmented by the Children's Health Insurance Program (CHIP) and the el-

igibility for children was expanded beyond the existing limits of the Medicaid program, most states required children to demonstrate being uninsured for certain period of time before being eligible. In general the waiting time varies from 3 to 6 months and it applies to children with family income above 150% of the poverty line (Lo Sasso and Buchmueller, 2004). Second, dropping a private insurance in favor of Medicaid would increase family income, affecting negatively health outcomes through food consumption effects. For instance, families may increase the consumption of "junk" food, which negatively affects children's health.

1.7 Conclusion

In this paper I analyze the effects of Medicaid on children's health care utilization and health outcomes. I estimate the causal effects of Medicaid taking advantage of Medicaid eligibility rule that generates a discontinuity in the probability of participating in Medicaid. In my analysis I account for potential heterogeneous effects of Medicaid on the health of children at different income eligibility levels, which is possible due to the variability of eligibility thresholds across states, time, and age groups.

My results highlight the importance of disaggregating the effects of Medicaid depending on the threshold level for a better understanding of the effects of the program. Indeed, my findings indicate that Medicaid induces a higher utilization of preventive medical care for the group of children at low threshold levels (below 185% of the poverty line) while it does not produce any significant change for the group of children at high threshold levels (between 185% and 250% of the poverty line).[36] The results also indicate that in the medium run −between 1 and 2 years after being eligible− Medicaid is more likely to have some persistent negative effects on some health outcomes, both at low and high threhold levels, although these effects vanishes afterwardws.

I proposed possible channels to explain the negative impact of Medicaid on children's health outcomes in the medium run which are consistent with the findings. First, a "perception" effect triggered by the increase in the contacts with physicians for preventive checkups. This explanation is more appealing for results obtained at low threshold levels. Second, a "quality" effect, as a consequence of Medicaid crowding out private insurance. The quality channel may be more suitable to explain the negative impact of Medicaid at high threshold levels, where the marginal groups is composed of children with higer family income levels. The quality channel explanation states that targeting higher income families with Medicaid may induce a crowding out effect and, although it might not affect health care utilization, it

[36] I cannot draw any conclusions for the threholds higher that 250% the poverty line, because the required condition to apply the fuzzy RD design, i.e., the probability of paticipating in Medicaid as a function of family income changes discontinuously at the threshold, is not satisfied.

might affect the quality of care a child can have access to. This switch may have undesirable health consequences for children as long as there are health care quality differences between Medicaid and private insurances.[37] Finally, I also discussed two alternative mechanisms (waiting times and consumption effects) through which the crowding out effect generated by Medicaid may negatively impact children's health outcomes.

These findings can provide a guide for improving the design and targeting of Medicaid. Medicaid is an effective tool to improve health care access at low threshold levels (hence at lower family income levels), but not at higher threshold levels. In terms of health, Medicaid seems to have null or negative effects on the dimensions analyzed in this paper. Given that I am analyzing the effects at the threshold, finding null effects may indicate that thresholds are optimally set, i.e., they are set such that the marginal beneficiaries obtain gains equal to zero.[38] The finding that Medicaid has some negative impact on children's probability of being in excellent health or that it increases the probability of being obese may be an indication of the potential consequences of the crowding effect. If quality differentials explains the negative impact on health, then there could be room for Medicaid quality improvements, without involving budgetary changes, that may help to reduce the negative unintended effects of Medicaid on higher income children. For example, better monitoring of the simple practices physicians treating Medicaid patients should follow may lead to better children's health outcomes. Particularly, improving the percentage of physicians that document the BMI and give counseling for nutrition and physical activity may be a cost-effective way to reduce the incidence of obesity on Medicaid eligible children. However, there is still a need for further research that tries to disentangle precisely the the mechanisms generating the observed negative effects.

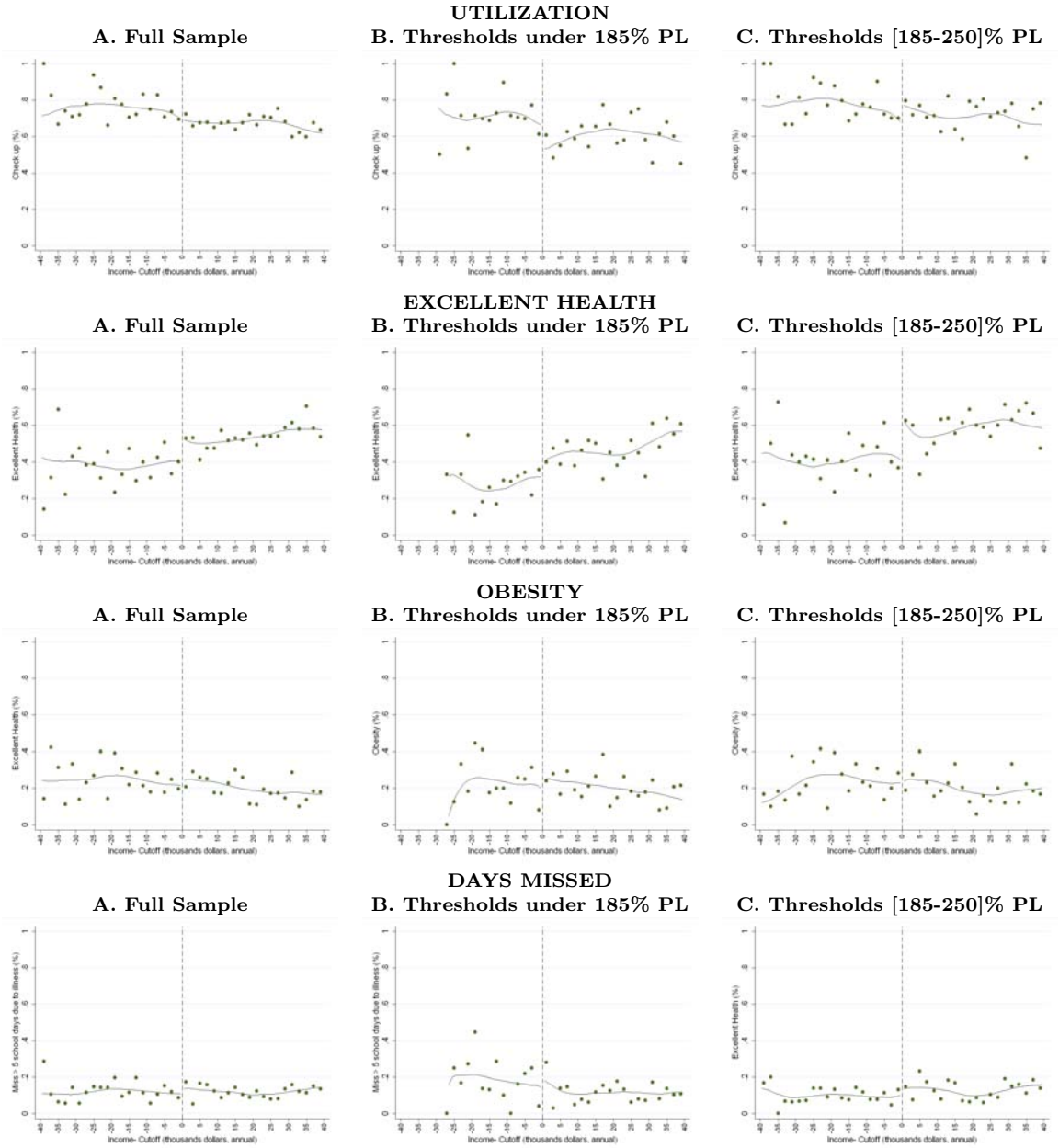
[37] Even when Medicaid also induces a crowding out effect at low thresholds, where the marginal affected group has lower levels of family income, it may not have an unintended effect on children's health. The reason is that if insurance quality is a normal good, then this group is more likely to buy, in the absence of Medicaid, low quality private insurances. Hence, at low thresholds switching into Medicaid is more likely to imply an increase in the quality of care.

[38] In this sense, states with low thresholds would obtain larger improvements in preventive medical care utilization by expanding their thresholds, because they are not exhausting all potential gains.

1.8 Appendix

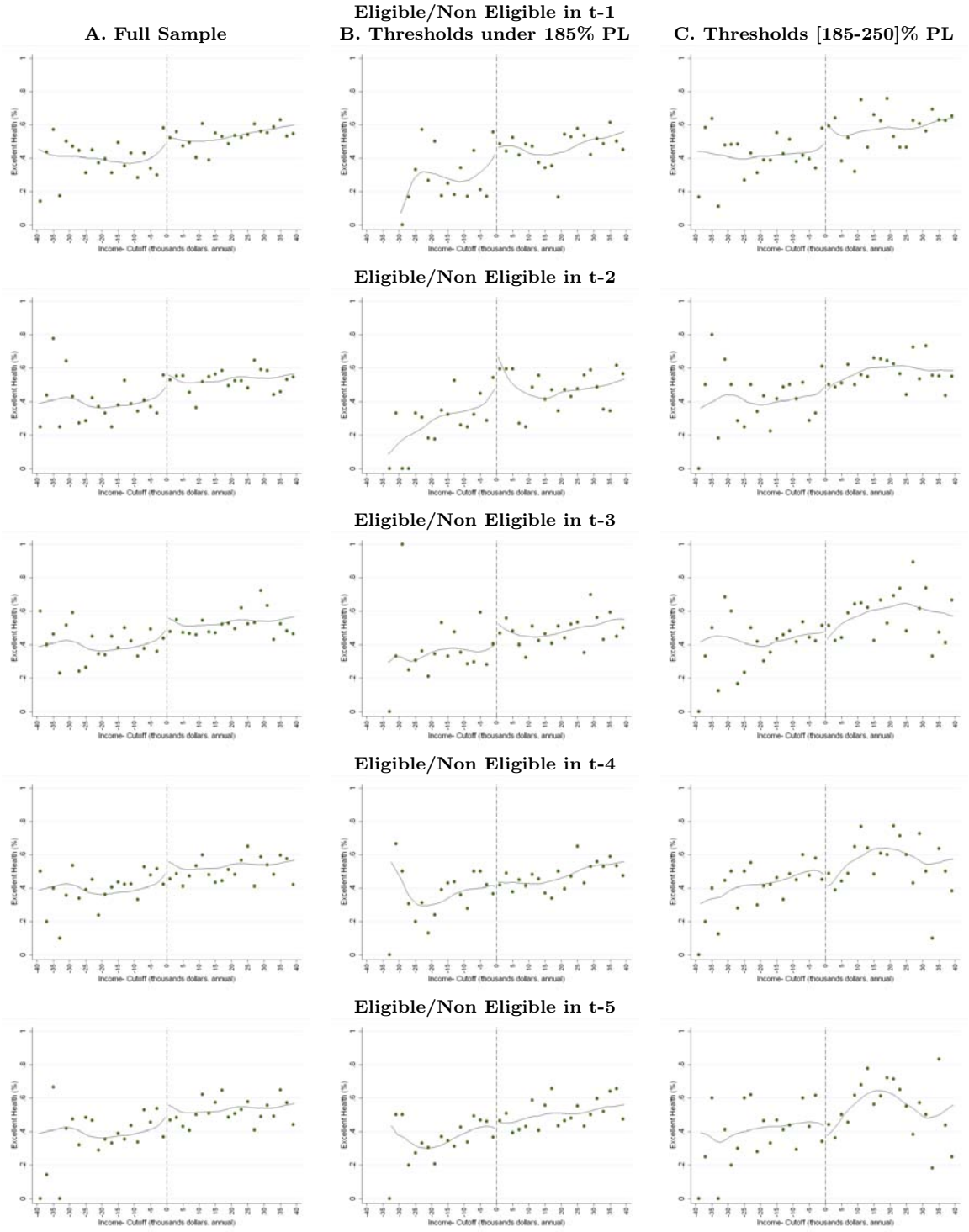
1.8.1 Graphs

Figure 1.4: Outcome Variables (contemporaneous eligibility)



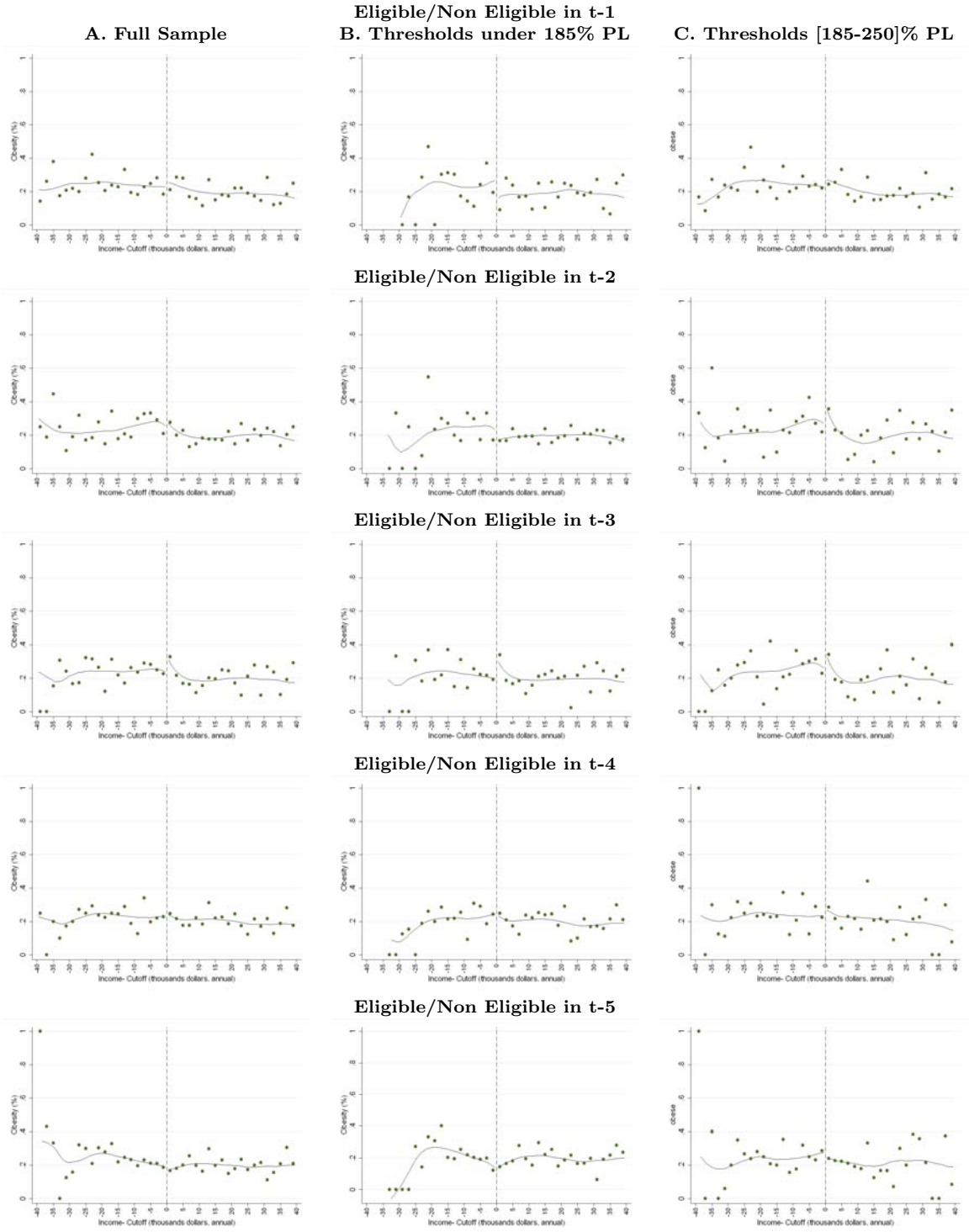
Solid lines are predictions from local linear regressions using triangle kernel, with the raw data.

Figure 1.5: Excellent Health (lagged eligibility)



Solid lines are predictions from local linear regressions using triangle kernel, with the raw data.

Figure 1.6: Obesity (lagged eligibility)



Solid lines are predictions from local linear regressions using triangle kernel, with the raw data.

1.8.2 Robustness Analysis of the Discontinuity in the Probability of Participating in Medicaid

In this appendix I perform a robustness analysis to show that the probability of participating in Medicaid as a function of family income is discontinuous at the eligibility threshold. I use the sample that considers family income and Medicaid participation for the whole period 1991 and 2007. Table 1.13 shows the estimated jump for different parametric specifications, confirming the pattern of Section 1.5.2.

Given that almost all states have thresholds set below 185% and between 185% and 250% at least once during the period 1991-2007, as it is shown in Table 1.14, I can extrapolate these results and say that on average children in higer income families are less likely to participate in Medicaid.

1.8.2.1 Placebo test for discontinuity

Figure 1.7: Discontinuity in the probability of participation. Placebo tests.

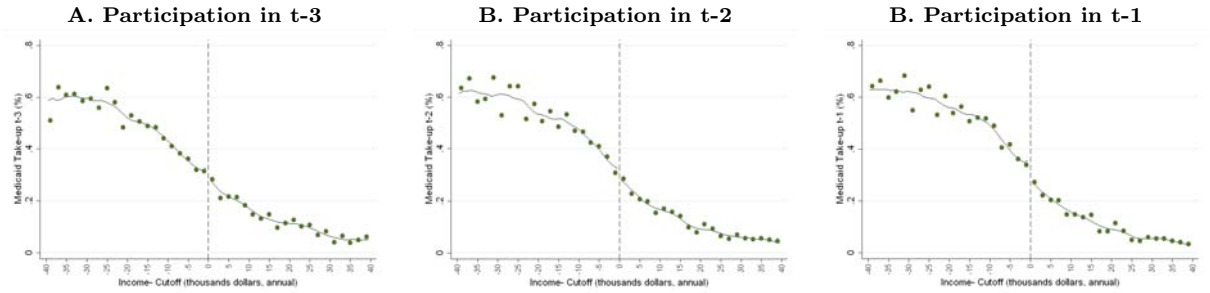


Table 1.13: Participation Equation. "Jump" at the threshold. Period 1991-2007.

	Bandwidth (thousands dollars)				
	±50	±30	±20	±15	±2
A. Full sample					
Polynomial Order					
One	0.208***	0.199***	0.158***	0.093***	0.074***
	(0.015)	(0.015)	(0.016)	(0.015)	(0.028)
Two	0.126***	0.104***	0.083***	0.063***	0.075***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.028)
Three	0.121***	0.086***	0.063***	0.057***	0.070***
	(0.014)	(0.014)	(0.015)	(0.015v	(0.028)
Four	0.112***	0.086***	0.064***	0.056***	0.073***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.028)
B. Model Interacted					
Polynomial Order					
One					
$Eli_t \times 1\{T < 185\}$	0.279***	0.275***	0.235***	0.151***	0.117***
	(0.021)	(0.021)	(0.022)	(0.023)	(0.042)
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.188***	0.178***	0.106***	0.095***	0.059
	(0.022)	(0.023)	(0.022)	(0.024)	(0.041)
$Eli_t \times 1\{T > 250\}$	0.022	0.051	0.029	0.053	-0.048
	(0.037)	(0.035)	(0.040)	(0.047)	(0.081)
Two					
$Eli_t \times 1\{T < 185\}$	0.178***	0.157***	0.141***	0.113***	0.114***
	(0.021)	(0.022)	(0.023)	(0.024)	(0.042)
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.115***	0.098***	0.073***	0.073***	0.056
	(0.020)	(0.021)	(0.023)	(0.023)	(0.042)
$Eli_t \times 1\{T > 250\}$	0.036	0.041	0.013	0.039	-0.054
	(0.039)	(0.039)	(0.039)	(0.046)	(0.081)
Three					
$Eli_t \times 1\{T < 185\}$	0.174***	0.128***	0.105***	0.096***	0.121***
	(0.020)	(0.022)	(0.024)	(0.025)	(0.042)
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.112***	0.076***	0.068***	0.063***	0.056
	(0.020)	(0.022)	(0.023)	(0.023)	(0.042)
$Eli_t \times 1\{T > 250\}$	0.053	0.031	0.017	0.037	-0.070
	(0.038)	(0.038)	(0.040)	(0.046)	(0.089)
Four					
$Eli_t \times 1\{185 \leq T \leq 250\}$	0.163***	0.130***	0.107***	0.100***	0.129***
	(0.021)	(0.022)	(0.024)	(0.025)	(0.042)
$Eli_t \times 1\{T > 250\}$	0.109***	0.087***	0.068***	0.064***	0.055
	(0.020)	(0.022)	(0.023)	(0.024)	(0.042)
$Eli_t \times 1\{T < 185\}$	0.058	0.036	0.017	0.039	-0.069
	(0.039)	(0.038)	(0.041)	(0.046)	(0.090)
N	22,701	17,857	13,391	10,411	1,426

Notes: **Panel A:** Each entry comes from a separate linear probability model $M_{i,t} = \pi_0 + \pi_1 Eli_{it} + k_{1g}(z_{it}; \alpha_{1g}) + u_{it}$. All regressions include a polynomial of the indicated order of log income, age, and family size; year and state dummies. **Panel B:** Each entry comes from a separate linear probability model $M_{it} = \gamma_0 + \sum_{j=1}^{2} \gamma_j T_{j,it} + \sum_{j=0}^{2} \pi_j Eli_{j,it} + k_{0g}(z_{it}; \alpha_{0g}) + \sum_{j=1}^{2} k_{jg}(z_{it}; \alpha_{jg}) \times T_{j,it} + u_{it}$. All regressions include a polynomial of the indicated order of the log income, age, and family size; year and state dummies. Robust standard errors (in parenthesis) are clustered at the family level. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated.

Table 1.14: Eligibility Thresholds by State

	The state has at least once, during the period 1991-2007, a threshold:		
State	under 185 % the FPL	[185,250] % the FPL	over 250 % the FPL
Alabama	Yes	Yes	No
Alaska	Yes	Yes	No
Arizona	Yes	Yes	No
Arkansas	Yes	Yes	No
California	Yes	Yes	Yes
Colorado	Yes	Yes	No
Connecticut	Yes	Yes	Yes
Delaware	Yes	Yes	No
District of Columbia	Yes	Yes	No
Florida	Yes	Yes	No
Georgia	Yes	Yes	No
Hawaii	Yes	Yes	Yes
Idaho	Yes	Yes	No
Illinois	Yes	Yes	No
Indiana	Yes	Yes	No
Iowa	Yes	Yes	No
Kansas	Yes	Yes	No
Kentucky	Yes	Yes	No
Louisiana	Yes	Yes	No
Maine	Yes	Yes	No
Maryland	Yes	Yes	Yes
Massachusetts	Yes	Yes	No
Michigan	Yes	Yes	No
Minnesota	Yes	Yes	Yes
Mississippi	Yes	Yes	No
Missouri	Yes	Yes	Yes
Montana	Yes	No	No
Nebraska	Yes	Yes	No
Nevada	Yes	Yes	No
New Hampshire	Yes	Yes	Yes
New Jersey	Yes	Yes	Yes
New Mexico	Yes	Yes	No
New York	Yes	Yes	No
North Carolina	Yes	Yes	No
North Dakota	Yes	No	No
Ohio	Yes	Yes	No
Oklahoma	Yes	Yes	No
Oregon	Yes	Yes	No
Pennsylvania	Yes	Yes	No
Rhode Island	Yes	Yes	No
South Carolina	Yes	Yes	No
South Dakota	Yes	Yes	No
Tennessee	Yes	Yes	Yes
Texas	Yes	Yes	No
Utah	Yes	Yes	No
Vermont	Yes	Yes	Yes
Virginia	Yes	Yes	No
Washington	Yes	Yes	No
West Virginia	Yes	Yes	No
Wisconsin	Yes	Yes	No
Wyoming	Yes	Yes	No
	51	49	10

1.8.3 Lagged effects

Table 1.15: Lagged cumulative intention to treat effects of Medicaid on Number of School Days Missed due to Illness. Children between 5 and 18 years old. Full sample (all thresholds pooled).

	Bandwidth (thousands dollars)					
Polynomial Order	±30	±20	±15	±30	±20	±15
	Eligible One Year Before (θ_1)					
One						
All Thresholds	-0.032	0.188	0.090	0.154	0.210	0.348
	(0.379)	(0.322)	(0.355)	(0.339)	(0.302)	(0.358)
Two						
All Thresholds	-0.204	0.017	0.515	0.138	0.285	0.293
	(0.431)	(0.331)	(0.379)	(0.450)	(0.386)	(0.444)
Polynomial Order	**Eligible Two Years Before (θ_2)**					
One						
All Thresholds	-0.341	-0.523	-0.473	-0.176	-0.323	-0.415
	(0.321)	(0.386)	(0.393)	(0.317)	(0.377)	(0.375)
Two						
All Thresholds	-0.384	-0.387	-0.264	-0.408	-0.422	-0.129
	(0.359)	(0.410)	(0.397)	(0.500)	(0.579)	(0.530)
Polynomial Order	**Eligible Three Years Before (θ_3)**					
One						
All Thresholds	-0.032	-0.041	-0.063	-0.049	-0.041	-0.148
	(0.040)	(0.044)	(0.051)	(0.348)	(0.418)	(0.400)
Two						
All Thresholds	-0.030	-0.053	-0.022	-0.054	0.017	0.009
	(0.047)	(0.051)	(0.062)	(0.562)	(0.616)	(0.544)
Polynomial Order	**Eligible Four Years Before (θ_4)**					
One						
All Thresholds	-0.300	-0.495	-0.150	-0.256	0.140	0.162
	(0.365)	(0.467)	(0.424)	(0.394)	(0.448)	(0.466)
Two						
All Thresholds	-0.931	-0.969	-0.177	-0.084	-0.114	-0.097
	(0.573)	(0.753)	(0.509)	(0.623)	(0.737)	(0.575)
Polynomial Order	**Eligible five Years Before (θ_5)**					
One						
All Thresholds	-0.089	-0.228	-0.107	0.045	-0.112	-0.102
	(0.277)	(0.347)	(0.394)	(0.347)	(0.388)	(0.433)
Two						
All Thresholds	-0.280	-0.441	-0.228	-0.314	-0.165	0.027
	(0.370)	(0.444)	(0.446)	(0.487)	(0.537)	(0.544)
Different Polynomial at each side	N	N	N	Y	Y	Y

Notes: Robust standard errors (in parenthesis) are clustered at the family level. All regressions are linear probability models and all include a polynomial of log income (rescaled to equal zero at the threshold); age and family size and its squares; year and state dummies. In each column the sample is restricted to observations with family income levels that falls within the bandwidth indicated. The lagged intention to treat estimates in each column come from the following model: $y_{it} = \alpha + \theta_\tau Eli_{i,t-\tau} + f_g(z_{i,t-\tau}) + f_g(z_{i,t-\tau}) \times Eli_{i,t-\tau} + u_{it}$. The first three columns do not allow for different polynomials at each side of the cutoff.

Chapter 2

Spillovers of Health Education at School on Parents' Physical Activity

(joint work with M. Lucila Berniell and Nieves Valdes)

Abstract

To prevent modern health conditions like obesity, cancer, cardiovascular illness, and diabetes, which have reached epidemic-like proportions in recent decades, many health experts argue students should receive Health Education (HED) at school. Although this type of education aims mainly to improve children's health profiles, it might affect other family members as well. This paper exploits state HED reforms as quasi-natural experiments to estimate the causal impact of HED received by children on their parents' physical activity. We use data from the Panel Study of Income Dynamics (PSID) for the period 1999-2005 merged with data on state HED reforms from the National Association of State Boards of Education (NASBE) Health Policy Database, and the 2000 and 2006 School Health Policies and Programs Study (SHPPS). To identify the spillover effects of HED requirements on parents' behavior we use a "differences-in-differences-in-differences" (DDD) methodology in which we allow for different types of treatments. We find a positive effect of HED reforms at the elementary school on the probability of parents doing light physical activity. Introducing major changes in HED increases the probability of fathers engaging in physical activity by 20 percentage points, although the probability of mothers being physically active did not seem to be affected. We find evidence of two channels that may drive these spillovers. We conclude that the gender specialization of parents in childcare activities, as well as information sharing between children and parents, may play a role in generating these indirect effects and in turn, in shaping healthy lifestyles within the household.

2.1 Introduction

Non-communicable diseases such as obesity, cancer, cardiovascular conditions, and diabetes have reached epidemic-like proportions in recent decades. Physical inactivity is one of the most important risk factors for these diseases (WHO, 2003). As a result, prevention increasingly involves changes in lifestyles, such as introducing the practice of regular physical activity in order to reduce risk factors (Kenkel, 2000). In the US, physically active individuals save an estimated US$ 500 per year in health care costs according to 1998 data (WHO, 2003).

Interactions within the family may crucially affect the "production" of healthy lifestyles. As Kenkel (2000) points out, the family is often identified as the unit of production of preventive practices. Previous literature on intra-household health decisions has focused on the interactions between spouses.[1] As well, the literature on intergenerational transmission of characteristics such as health, ability, education or income, has focused on the effects that parents' decisions can have on children's behaviors and outcomes.[2] However, little research has been done to evaluate the impact of children on parents' decisions, in particular on healthy lifestyle choices.

Schools can play a fundamental role in providing children with information about healthy lifestyles and health decisions, which may complement what they learn at home. At school, the knowledge about health is transferred to children through the implementation of specific curricular modules, often known as Health Education (HED).[3] Although HED is likely to affect children's health behaviors, it may be the case that parents are also affected by the education about preventive health care that their children acquire at school.[4] Moreover, the indirect effect of HED on parents may in turn enhance the effectiveness of HED delivered in the school setting in changing children's health behaviors.

The first goal of this paper is to assess the existence of spillover effects of Health Education received by children at school on their parents.[5] We exploit the quasi-experiment provided by the changes in the state-level HED requirements in elementary schools implemented between the school years 1999/2000 and 2005/2006 in the US to quantify the effects of

[1] For instance, see Clark and Etile (2006) on spousal correlation of smoking behavior.

[2] There are numerous studies quantifying the role of intergenerational transmission of parents' characteristics and behaviors on children's outcomes (Currie, 2009).

[3] According to the Centers for Disease Control and Prevention (CDC) "*Health Education is a planned, sequential, and developmentally appropriate instruction about Health Education designed to protect, promote, and enhance health literacy, attitudes, skills, and well-being*" (Kann, Telljohann, and Wooley, 2007).

[4] As stated by WHO (1999), there are several reasons for promoting healthy behaviors through schools. Schools are an efficient way to reach school-age children and their families in an organized way and students spend a great portion of their time in schools, where education and health programs can reach them at influential stages in their lives.

[5] For instance, providing physical education at school has proven to be an effective way to improve healthy habits in children (Cawley, Meyerhoefer, and Newhouse, 2007).

these programs on parents' physical activity.[6] Thus, the focus is on a policy that does not imply any transfer of resources to children -the targeted individuals- but instead provides them with new information. A second goal of this paper is to discuss the plausible channels through which children receiving HED at schools may affect the probability that their parents engage in physical activity.

To identify the spillover effects of HED policies, we use a "differences-in-differences-in-differences" (DDD) strategy, exploiting not only the time series and cross-state variation, but also within-state variation. The time dimension allows us to include year effects in order to capture national trends in physical activity. The variation across states allows for controlling for systematic differences in physical activity between people living in states that change their HED policies and people living in states that do not change their HED policies. The variation within states makes possible controlling for state-specific time trends that can be correlated with the change in HED policies. We are able to exploit the third difference because within each state there are individuals who were exposed to the treatment and others who were not. We show in Section 2.3 that there are remarkable differences in the pre-treatment trends in the outcomes of experimental versus non-experimental states, indicating that the implementation of HED policies is correlated with the behavior of the outcome of interest, which makes the use of a DDD estimator crucial here. The data we use is from the Panel Study of Income Dynamics (PSID) for the period 1999-2005, merged with data on state HED reforms from the State School Healthy Policy Database of the National Association of State Boards of Education (NASBE), and the 2000 and 2006 surveys of the School Health Policies and Programs Study (SHPPS).

Our results show evidence of a positive effect of HED received by children in elementary schools on their fathers' probability of engaging in physical activity. Introducing major reforms in HED in elementary schools makes a father exposed to this policy 20 percentage points more likely to be physically active than a comparable father not affected by the policy. We do not find evidence that the policy under analysis affects the decision of mothers to engage in physical activity.

We explore the channels behind these results, and find two non-exclusive explanations. First, we argue that a "role model" channel may explain the differential impact according to parent gender. In effect, the roles that mothers and fathers play for their children in the activities they usually do together are important for this result. Parents usually spend more time with their children doing gendered activities, such as physical activity in the case of fathers. Therefore, the promotion of healthy behaviors at school is more likely to have an effect on the behavior of fathers than that of mothers. Second, we find evidence consistent with an "information sharing" channel. We analyze the differential impact of HED reforms on individuals with low and high education and income levels and find greater

[6] Further details on these policy reforms can be found in Section 2.2.

effect on individuals with lower education and income levels. The existence of spillovers of HED on parents' lifestyles indicates that the interaction between children and parents plays a role in the formation of healthy lifestyles within the household, which must be taken into account to properly design policy interventions aimed at increasing the adoption of healthy lifestyles in a given community.

We perform a number of robustness checks that support the causality of the link between HED received by children in elementary schools and the probability of their fathers engaging in physical activity. First, we show that HED reforms do not affect outcomes that are not related to health behaviors, such as labor force participation. Second, we perform a "placebo" test on adults that were not exposed to the potential indirect effect of HED. The test shows that the placebo treatment group is not affected by the HED reforms, indicating that our results are not driven by other shocks contemporaneous to HED changes that systematically affected parents in the treatment group. Finally, we show that our results are also robust to alternative definitions of the control group.

This work is related to two strands of literature. First, it is related to the literature on policy evaluation that focuses on measuring the spillover effects of policy interventions on non-targeted individuals, also known as Indirect Treatment Effects (ITE). We focus on spillovers of a program targeting children on parents' behavior. We know of two interventions explicitly designed to have school-age children affecting their families and other community members health behaviors. Harre and Coveney (2000) evaluate two pilot studies implemented in a New Zeland school that taught children aged 7-11 years about burns and scalds hazards, and encouraged changes to the home environment and family practices through a take-home exercise. The intervention was designed to have an impact on the safety knowledge and behavior of primary school children and their parents. Nandha and Krishnamoorthy (2007) describe the role and effectiveness of school-based HED for social mobilization to promote the use of a fortified salt in an Indian district where lymphatic filariasis is endemic. HED through classroom sessions was the main motivational strategy used in this intervention that targeted community members to receive the message through children. Regrettably, both case studies lack the ability to state causality since the interventions were not randomly assigned and affected few individuals. There are a small number of works in the economic literature assessing the existence of spillovers on non-targeted individuals within the household that present reliable results by using neat identification methodologies. One exception is Bhattacharya, Currie, and Haider (2006), who analyze the effects of the School Breakfast Program (SBP) in the US on not only targeted children but also on adult (non-targeted) family members. They find that the SBP improves the quality of diets even for family members who were not directly exposed to the program.[7]

[7] Jacoby (2002) and Shi (2008) also analyze the effects of policies directed at children on non-eligible members of the household. They do not find evidence of family spillover effects. Jacoby (2002) analyzes

The explanation for family spillover effects in this work is that the particular program reduces family budgetary constraints, freeing resources that may be redirected towards other household members. In contrast, we explore family spillovers occurring for non-budgetary reasons. There are also some works in this literature evaluating external effects at the community level instead of the family level. Some examples are Angelucci and Giorgi (2009), Lalive and Cattaneo (2006), and Miguel and Kremer (2004).[8]

The second strand of literature related to our work consists of recent research evaluating the direct impact of particular aspects of health education at the school level on students' health outcomes and behaviors. Cawley, Meyerhoefer, and Newhouse (2007) find positive effects of physical education requirements on the amount of time high school students engage in physical exercise, although they do not find any impact on Body Mass Index (BMI) or the probability of students being overweight. Also, McGeary (2009) assesses the effects of state-level nutrition education program funding on the BMI, the probability of obesity, and the probability of above normal weight.[9] Her results suggest that this funding is associated with reductions in BMI and in the probability of an individual having an above-normal BMI. Kahn, Ramsey, Brownson, Heath, Howze, Powell, Stone, Rajab, and Corso (2002), Salmon, Booth, Phongsavan, Murphy, and Timperio (2007) and van Sluijs, McMinn, and Griffin (2007) summarize the results of several interventions aimed at evaluating the effectiveness of HED programs in changing children's physical activity. The three articles agree that the interventions reviewed provided insufficient evidence to assess the effectiveness of classroom-based HED and family-based social support interventions in increasing levels of physical activity or improving fitness because of inconsistent results among studies and various limitations in the studies design.[10]

---

the impact of a school meals program in the Philippines on caloric intake of targeted and non-targeted individuals in the family, whereas Shi (2008) studies resource reallocation in the household after a child receives a subsidy to cover school fees in rural China. These two papers find evidence of intra-household flypaper effects, that is no sizable reallocation of resources after children receive subsidies.

[8] Angelucci and Giorgi (2009) evaluate the spillover effects of an aid program (PROGRESA) on entire local economies (villages) where the program was implemented. Lalive and Cattaneo (2006) find that PROGRESA significantly increases school enrollment among non-eligible families in the villages and that this rise is driven by a peer effect. Using evidence from a randomized experiment, Miguel and Kremer (2004) show that a deworming program substantially improved health and school participation among untreated children in both treatment schools and neighboring schools.

[9] This funding is allocated to public-school systems, public-health clinics, as well as public-service announcements and advertisements. McGeary's analysis goes beyond the effects of education at school, and therefore she computes the estimates for the entire population in each state.

[10] Among the several limitations in the studies the authors single out the lack of information on the randomization procedure, short duration of follow-ups, lack of precision of the physical activity outcome measures, and small sample sizes.

2.2 Health Education Policies in the US

In the 1970s and 80s, research studies showed that healthy kids did better in school and scored higher on achievement tests. As a consequence, some states started to develop and implement HED programs in public schools. In the 1990s, many educators called for the creation of a set of national health education standards that states could use as a template. In 1995, the National Committee for Health Education Standards created national health education standards with K-12 benchmarks covering several content areas of health. In 1998, the Congress urged the Centers for Disease Control and Prevention (CDC) to "expand its support of coordinated health education programs in schools" (Wyatt and Novak, 2000). As Kahn, Ramsey, Brownson, Heath, Howze, Powell, Stone, Rajab, and Corso (2002) explain, "*HED classes that provide information and skills related to decision making are usually multicomponent, with the curriculum typically addressing physical activity, nutrition, smoking, and cardiovascular disease. HED classes, taught in elementary, middle, or high schools, are designed to effect behavior change through personal and behavioral factors that provide students with the skills they need for rational decision making*".

State HED programs are typically characterized by two dimensions. The first is the health education curricula indicating the health related topics schools are *required* to teach. Panel A of Table 2.1 lists the topics included as potential HED requirements. We focus in these five topics because all of them may affect the knowledge about the benefits of being physically active.[11]

The second dimension is specific regulations to guarantee and strengthen the effective and coordinated implementation of health education in schools. We broadly refer to these regulations as *enforcements*. Panel B of Table 2.1 describes the three specific state requirements enforcing HED we focus on.[12]

In the period 1994 and 1999 school health policies at the state level generally remained unchanged, but important changes were detected between 1999 and 2005.[13] During this period, states either implemented HED programs for the first time or expanded one or both dimensions of pre-existing programs.

2.2.1 Databases for HED programs: NASBE and SHPPS

The information we use to define which states have HED programs and the degree of development of such programs -i.e., which topics were required and which enforcements were mandatory at different points in time- comes from two complementary sources: the

[11] Table 2.10 in the Appendix shows other topics that could potentially be included in an elementary school HED curriculum, but we do not take them into consideration because they are more related to sex education.

[12] The full list of potential requirements is shown in Table 2.10 in the Appendix.

[13] See Kann, Brener, and Allensworth (2001) and Kann, Telljohann, and Wooley (2007) for more details on these changes in policies.

Table 2.1: HED Programs

A) Curricula: Topics covered
1) Alcohol- or Other Drug-Use Prevention
2) Emotional and Mental Health
3) Nutrition and Dietary Behavior
4) Physical Activity and Fitness
5) Tobacco-Use Prevention

B) Enforcements
1) State requires districts or schools to follow national or state health education standards or guidelines
2) State requires students in elementary school to be tested on health topics
3) State requires each school to have a HED coordinator

NASBE State School Health Policy Database and the School Health Policies and Programs Study (SHPPS).

The NASBE Database is a comprehensive set of laws and policies of all states in the US on more than 40 school health policies. It began in 1998 and is maintained with support from the Division of Adolescent and School Health (DASH) of the CDC. The database contains brief descriptions of laws, legal codes, rules, regulations, administrative orders, mandates, standards, resolutions, and other written means of exercising authority. While authoritative binding policies are the primary focus of the database, it also includes guidance documents and other non-binding materials that provide a detailed picture of a state's school health policies and activities.

The NASBE Database was designed to build upon the SHPPS, conducted by the CDC every 6 years since 1994. SHPPS is a nationwide survey that gathers detailed and comparable information about the characteristics of HED programs at the state level across elementary, middle, and high schools.[14] While SHPPS collects state policy information by means of survey questionnaires that are completed by state education agency personnel, the NASBE Database provides the legal support for the policies reported in SHPPS.

Using the information provided by both sources we classified each state as either an *"Experimental State"*, if the state changed the HED program between 1999 and 2005, or as a *"Non-Experimental State"*, if no changes were introduced in the state HED program during the period. Tables 2.11 and 2.12 in the Appendix give a detailed description of HED programs in all states in 1999 and 2005.

[14] SHPPS also gathers information about health-related programs at the district, school, and classroom levels. SHPPS analyzes eight components, one of which is the HED component. The remaining seven components are physical education and activity, health services, mental health and social services, nutrition services, a healthy and safe school environment, and faculty and staff health promotion.

2.3 Identification Strategy and Data

Our goal is to identify the spillover effects of elementary school HED policies implemented in certain states (the "experimental states") on the behavior of parents of elementary school-age children (the treatment group). Identifying this effect requires, as stated in Gruber (1994), controlling for any systematic shocks to the parents' outcome behavior in the experimental states that are correlated with, but not due to, changes in HED policies. To do so, we use a "differences-in-differences-in-differences" (DDD) approach that allows us to exploit the variation of HED policies across time (time dimension), across states (geographical dimension), and across different groups of individuals residing in the same state (individual dimension). That is, we compare the treatment individuals in experimental states to a set of control individuals in those same states and we measure the change in the treatments' relative outcome, relative to those of states that did not change HED policies. The identifying assumption requires that there is no contemporaneous shock affecting the relative outcome of the treatment group in the same state-year as the change in the HED policy.

We use a DDD identification strategy instead of the more commonly used "difference-in-differences" (DD) because it does not require the common trend assumption for treatment and control groups. We consider that this assumption will most likely be violated given the characteristics of the policy we are analyzing. In particular, the DD estimator of the spillover effects of HED policies on parents will be biased if the states that increased their HED requirements between 1999 and 2005 were those where health indicators were deteriorating more rapidly. To explore this possibility, we looked at health indicators of the population of adults with children below 18 years of age for pre-treatment periods (1994-1998), using data from the Behavioral Risk Factor Surveillance System (BRFSS).[15] As shown in Table 2.2, the proportion of individuals at risk because of overweight or obesity has increased more rapidly in experimental states than in non-experimental states. As well, between 1994 and 1998 the proportion of individuals with sedentary lifestyles has increased more in experimental than in non-experimental states. Therefore, the different trends in the outcomes of experimental versus non-experimental states indicate that the implementation of HED policies is correlated with the evolution of the outcome of interest, which makes the use of a DDD estimator crucial here.

Formally, let y_{it}^1 be the outcome for individual i at time t if she/he is exposed to the treatment. The outcome for the same individual if not exposed to the policy is y_{it}^0. Consequently, the impact of the policy on individual i is $y_{it}^1 - y_{it}^0$. The average treatment effect across treated individuals is $\mathbb{E}(y_{it}^1 - y_{it}^0 | elem = 1, S = 1)$, where $elem = 1$ denotes individuals who

[15] Note that we made use of this other dataset to evaluate the pre-treatment trends because the PSID does not contain information on health issues for this period of time.

Table 2.2: Lack of common trends between experimental and non-experimental states.

Year	Obesity 1 (%)		Obesity 2 (%)		Sedentary lifestyle (%)	
	Non-exper. states	Exper. states	Non-exper. states	Exper. states	Non-exper. states	Exper. states
1994	33,1	32,8	28,3	28,4	59,5	56,9
	(22824)	(13693)	(22824)	(13693)	(22824)	(13693)
1996	35,7	35,6	30,7	31,5	59,0	59,4
	(24612)	(16470)	(24612)	(16470)	(24612)	(16470)
1998	36,9	39,8	32,4	34,9	57,0	59,1
	(29052)	(20767)	(29052)	(20767)	(29052)	(20767)
Var. % ('94-'98)	11,6%	21,3%	14,4%	22,7%	-4,2%	4,0%

Source: BRFSS 1994, 1996, and 1998. Sample sizes in parentheses. *Definitions:* **Obesity 1 (%)**: Percentage of population (with children under 18 years old) at risk for obesity (greater than 120% of weight for height percent median). **Obesity 2 (%)**: Percentage of population (with children under 18 years old) at risk for overweight based on BMI. At risk defined as: >27.8 for males and >27.3 for females. **Sedentary lifestyle (%)**: Percentage of population (with children under 18 years old) at risk for sedentary lifestyle (sedentary or irregular physical activity profile).

have elementary school-age children –the treatment group– and $S = 1$ denotes individuals who reside in a state where HED requirements changed between 1999 and 2005 –the experimental states–. The treated individual has both, $elem = 1$ and $S = 1$. In our setup, the methodological challenge is to obtain a way to estimate the missing counterfactual $\mathbb{E}(y_{it}^0|elem = 1, S = 1, \tau_t = 1)$, where τ_t is a dummy variable, equal to one in 2005.
The population under analysis includes adults who have children living with them. The specification for the outcome is

$$\begin{aligned} y_{it} = {} & \beta_0 + \beta_1\tau_t + \beta_2 elem_i + \beta_3 S_i \\ & + \beta_4(elem_i \times \tau_t) + \beta_5(S_i \times \tau_t) + \beta_6(elem_i \times S_i) \\ & + \beta_{7i}(\tau_t \times elem_i \times S_i) + u_{it}, \end{aligned} \tag{2.1}$$

where $i = 1...N$ indexes individuals, and $t = 0, 1$ indexes time (0=before the policy change, 1999; 1=after the policy change, 2005). As stated before, τ_t is a dummy variable, equal to one in 2005, so it captures a nationwide time trend in the outcome; $elem_i$ is a dummy variable that takes the value one if individual i has at least one child of elementary-school-age, reflecting a group fixed effect in the outcome; and S_i is a dummy variable equal to one if individual i resides in an experimental state, that is, a state where the HED policy has changed between 1999 and 2005, allowing for an experimental-state fixed effect in the outcome. Moreover, the outcome may present differential time trends: (1) between parents of elementary school-age children versus parents of children of other ages and (2) between individuals living in experimental states and those living in non-experimental states. $(elem_i \times \tau_t)$ and $(S_i \times \tau_t)$ are the group-trend and the experimental state-trend respectively.

Since parents of elementary school-age children in experimental states may have a different outcome than parents of children below and above elementary school age also living in experimental states, we include the group-state fixed effect captured by the interaction $(elem_i \times S_i)$. Finally, the triple interaction $(\tau_t \times elem_i \times S_i)$ is equal to one only for treated individuals in the after-policy-change time period: these are the parents of elementary school-age children residing in experimental states in 2005.
The treatment effect for individual i is β_{7i}, and the average treatment effect on the treated (ATT) is $\mathbb{E}(\beta_{7i}|elem_i = 1, S_i = 1)$. The ATT can be recovered by sequential differences, up to the unobserved temporary individual-specific shocks u_{it}, that is

$$\begin{aligned} ATT &= \mathbb{E}(\beta_{7i}|elem_i = 1, S_i = 1) \\ &= [\mathbb{E}(y_{i1} - y_{i0}|elem_i = 1, S_i = 1) - \mathbb{E}(y_{i1} - y_{i0}|elem_i = 1, S_i = 0)] \\ &\quad - [\mathbb{E}(y_{i1} - y_{i0}|elem_i = 0, S_i = 1) - \mathbb{E}(y_{i1} - y_{i0}|elem_i = 0, S_i = 0)], \end{aligned} \tag{2.2}$$

only if:

$$\begin{aligned} &\mathbb{E}(u_{i1} - u_{i0}|elem_i = 1, S_i = 1) - \mathbb{E}(u_{i1} - u_{i0}|elem_i = 1, S_i = 0) \\ &= \mathbb{E}(u_{i1} - u_{i0}|elem_i = 0, S_i = 1) - \mathbb{E}(u_{i1} - u_{i0}|elem_i = 0, S_i = 0). \end{aligned} \tag{2.3}$$

The assumption (2.3) will hold if the outcome of parents in the treatment group in experimental states relative to the outcome of the same group of parents in non-experimental states is affected in the same way by idiosyncratic temporary shocks as the relative outcome of parents in the non-treatment group in experimental and non-experimental states.
The sample analog of equation (2.2) is the DDD estimator of the ATT:

$$\begin{aligned} \widehat{ATT} = &(\bar{y}_1^{1,1} - \bar{y}_0^{1,1}) - (\bar{y}_1^{1,0} - \bar{y}_0^{1,0}) \\ &- [(\bar{y}_1^{0,1} - \bar{y}_0^{0,1}) - (\bar{y}_1^{0,0} - \bar{y}_0^{0,0})], \end{aligned} \tag{2.4}$$

where $\bar{y}_t^{elem,S}$ is the average of the estimated outcome among individuals in group $elem$, residing in states S, at time t.
We can derive the same estimator for the ATT by recovering the missing counterfactual $\mathbb{E}(y_{it}^0|elem_i = 1, S_i = 1, \tau_t = 1)$, and rewriting the ATT as a function of unobserved counterfactuals using equation (2.1).

2.3.1 Database

We analyze the impact of HED policies on the behavior of adults who have children attending elementary school using data from two sources. The information on HED policies is obtained from the NASBE Database and the SHPPS, and the information on individuals

comes from the Panel Study of Income Dynamics (PSID).

The PSID is a nationally representative longitudinal survey of individuals in the US (men, women, and children) and the family units in which they reside. In 1999, the PSID has expanded the set of health-related questions for the heads of family units and spouses, gathering information on health status, health behaviors, health insurance, and health care expenditures. We concentrate on the indirect effect of HED policies on levels of physical activity, which is one of the health behaviors reported in this survey. The PSID also provides detailed information about family income, as well as family composition and demographic variables, including the ages of family members, race, marital status, employment status and education. The PSID covers all states.

We base our analysis on the PSID survey years 1999 and 2005, using 1999 as the pre-reform period. The DDD design we use to identify the effect of interest does not require the use of a panel, but the identification is improved by using longitudinal data. Even though we do not specify a model for panel data, in our final sample about 90% of the observations correspond to individuals in a panel.

Our final sample consists of 10,663 observations that include parents of children living with them, who participated in the 1999 and/or 2005 PSID. It is worth noting that for most of the individuals we also have her/his spouse or partner in the sample. Given the way in which the PSID is designed, for some of the individuals we also have another relative in the sample, for instance siblings. This feature of our data makes it important to estimate robust standard errors clustered at the family level.

Besides the PSID, there are other household and individual surveys containing information about health lifestyles. However, these surveys do not include all the variables we require to conduct our analysis for the years in which we can identify HED policy changes. The National Health Interview Survey (NHIS) and the National Health and Nutrition Examination Survey (NHANES) gather rich information about health, health behaviors, and socio-demographic characteristics. However, in both surveys the public-use data files do not include the state identifiers necessary to create HED reform variables at state level. Also, the Behavioral Risk Factor Surveillance System (BRFSS) has information about health behaviors and demographic variables, but its information on the age of children is incomplete.[16] Finally, the National Longitudinal Survey of Youth 1979 (NLSY79) recovers some information about health behaviors, but the information about adult's physical activity is not available for the years for which we can construct the policy reform variables.

[16] In the 1999 BRFSS survey, there are some available variables indicating the number of children younger than 5 years old, the number of children between 5 to 12 years old, and the number of children who are 13 through 17 years old within the household. Since November 2004 information about one randomly selected child, including age of the child, is available for some households. Hence, in those households where there is one child only, information about its age is available but there is missing information about the age of the other children in households with more than one child.

In the NASBE Database and in the SHPPS surveys we found that HED policies across states are highly heterogeneous, not only in terms of whether the state has implemented a HED program, but also regarding the scope and effectiveness in the implementation of such programs. Accordingly, we divided the non-experimental and experimental states into several groups. The non-experimental states are those states that did not change their HED policies between 1999 and 2005. We classified the non-experimental states into two groups: (1) States without HED programs in 1999 and 2005; (2) States with HED programs implemented by 1999, and without changes in 2005. We name groups (1) and (2) $S1$ and $S2$, respectively.

The experimental states are those that introduced any HED reforms between 1999 and 2005. There are three types of treatments (policies) that define three types of experimental states. Group $S3$ are states that, while having some topics in their HED curricula in 1999, did not introduce changes in those topics by 2005, but introduced some reforms in enforcements. Group $S4$ are states that, while having some topics required in 1999, increased the number of topics required by 2005, without introducing changes in enforcements. We consider that these two policies involve only minor changes in the already implemented HED programs, so in what follows we refer to these groups of states as "Moderate changes A" and "Moderate changes B", respectively. Finally, we include in the group $S5$ those states that for the first time introduced required topics at state level in their HED programs by 2005. We consider this policy to be a deep reform in HED, so we refer to group $S5$ as "Major changes". Some of the states introduced topics for the first time by 2005, while they did not make changes in enforcements, as were the cases of Arkansas and Florida. New Mexico and Wyoming introduced topics as mandatory by 2005, and simultaneously strengthened their HED policies by introducing new enforcements. A particular case is Texas, where all districts had a mandatory HED program in 1999 designed and implemented following district rules. It was not until 2005 that Texas implemented a coordinated HED program requiring all public schools in the state to have all topics in curriculum that followed national HED guidelines.

The information available in the NASBE database and SHPPS surveys regarding HED in the District of Columbia, Minnesota, and New Hampshire was not conclusive, so we could not classify these states and, consequently do not include them in our sample. We do not use states in our estimations for which the sample size was insufficient to control for temporal and group trends within the state.[17] Table 2.3 presents the aforementioned state

Table 2.3: States classification by changes in HED requirements between 1999 and 2005.

Group		Type of policy	Num. of states	Num. of Observations
Non-	$S1$	Does not have HED in 1999 and 2005	2	707
Experimental	$S2$	Existing HED in 1999 remains unchanged in 2005	18	6,417
	$S3$	Moderate changes A	5	1,099
Experimental	$S4$	Moderate changes B	5	1,156
	$S5$	Major changes	3	1,284
Total			33	10,663

Source: NASBE State School Health Policy Database, SHPSS surveys, and PSID database. The number of observations is the number of individuals in each group of states.

classifications and the sample sizes for the states included in our sample.[18]

We modify the specification in equation (2.1) to introduce the previous classification of states, and to allow for differential effects of the policy across different types of treatment

$$\begin{aligned} y_{it} = \beta_0 + \beta_1 \tau_t + \beta_2 elem_i + \sum_{k=2}^{5} \beta_{3,k} Sk_i \\ + \beta_4 (elem_i \times \tau_t) + \sum_{k=2}^{5} \beta_{5,k} (Sk_i \times \tau_t) + \sum_{k=2}^{5} \beta_{6,k} (elem_i \times Sk_i) \\ + \sum_{k=3}^{5} \beta_{7,k} (\tau_t \times elem_i \times Sk_i) + u_{it}, \end{aligned} \tag{2.5}$$

where, as before, $i = 1...N$ indexes individuals, and $t = 0, 1$ indexes time (0=before policy, 1999; 1=after policy, 2005), and now $k = 1, ..., 5$ indexes state groups.[19]

In our setting, treated individuals, those exposed to changes in HED policies, are adults who reside in an experimental state, and who have elementary school-age children (6-10). The PSID does not provide information on whether a child is attending elementary school. However, it provides information on the age of children, allowing us to determine if individuals have school-age children.[20]

The control group consists of individuals who were unaffected by changes in state HED requirements; it includes adults who have elementary school-age children (6-10) living in states

[17] States excluded from our database due to small sample size are Alaska, Delaware, Hawaii, Idaho, Maine, Montana, Nevada, New Mexico, North Dakota, Oklahoma, Rhode Island, South Dakota, Vermont, West Virginia, and Wyoming. To check that the exclusion of these states does not drive our results we estimate the effect of interest including the states with small sample size and the results are comparable.

[18] The complete list of states in each group, and the number of observations in each state are reported in Table 2.13 in the Appendix.

[19] $S1$ is the group of reference.

[20] Note that the dropout rate in elementary school is very low in the US. Therefore, by knowing the age of the children we are able to know whether the child is attending elementary education.

that did not change HED policies, that is, living in states that either did not implement HED policies or that, while having HED requirements in 1999, did not introduce any reform during the period. Furthermore, to control for possible correlation of state HED policies with unmeasured state trends in health and health behaviors, we use a sample of adults who have children living with them but not of elementary school age as a within-state comparison group. We group the non-treated individuals in three different control groups. We include in the Treatment-Non-Experimental group (Control 1) individuals with elementary school-age children residing in non-experimental states. The Control-Experimental group (Control 2) includes individuals with children not of elementary school-age residing in experimental states. Finally, in the Control-Non-Experimental group (Control 3) we include individuals with children above and below elementary school age residing in non-experimental states.

2.3.2 The outcome variable

Our outcome variable is light physical activity. PSID respondents are asked about their physical activity habits through two questions, the first about how often they do light physical activity and the second about the frequency of these activities (daily, weekly, monthly or annually). Based on these two questions we construct a variable indicating the number of times per week individuals do light physical activity. It is an ordinal variable with 49 different values, from 0 to 21. Its histogram is shown in Figure 2.1, according to which 15.4% of parents in the sample reported not doing any physical activity at all, while the remaining 84.6% reported engaging in light physical activity some number of times per week. Two well-differentiated mass points, at values 0 and 7, can be identified. As well, more than 12% of the observations lie in the interval (0,2), while another 34% are in the interval [2,7).

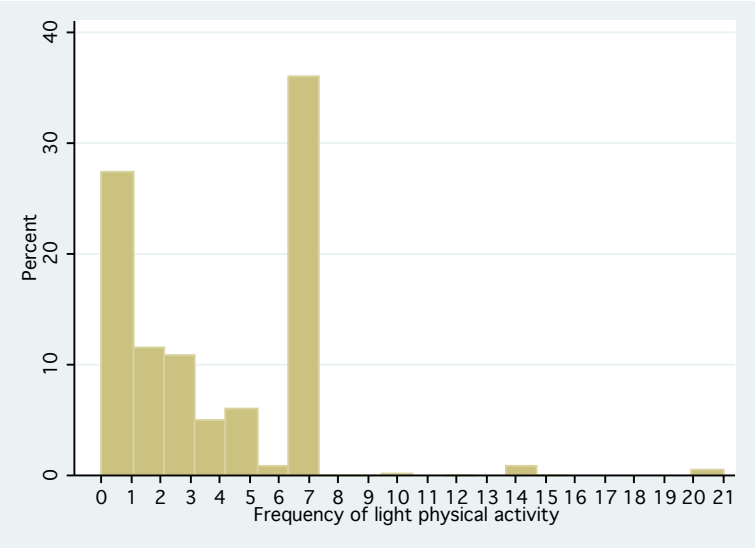
One limitation to using the number of times per week of light physical activity directly as our outcome variable is that there is no information in the PSID about the amount of time (minutes, hours) individuals spend each time they do physical activity. For example, in our database an individual that reports doing light physical activity three times per week is not necessarily doing more light physical activity than an individual that reports one session per week. This makes it very difficult to compare individuals who are physically active. To overcome this problem, we use as the outcome variable a binary variable that reflects whether an individual reports engaging in light physical activity at least once a week.

In what follows, the outcome variable is

$$y_i = \begin{cases} 1 & \text{if } i \text{ does light physical activity at least once a week,} \\ 0 & \text{otherwise.} \end{cases}$$

The two graphs in the left panel in Figure 2.2 show the proportion of physically active individuals by gender in 1999 and 2005 for the treated and control groups. We observe a

Figure 2.1: Histogram for the variable "frequency of light physical activity" (times per week).



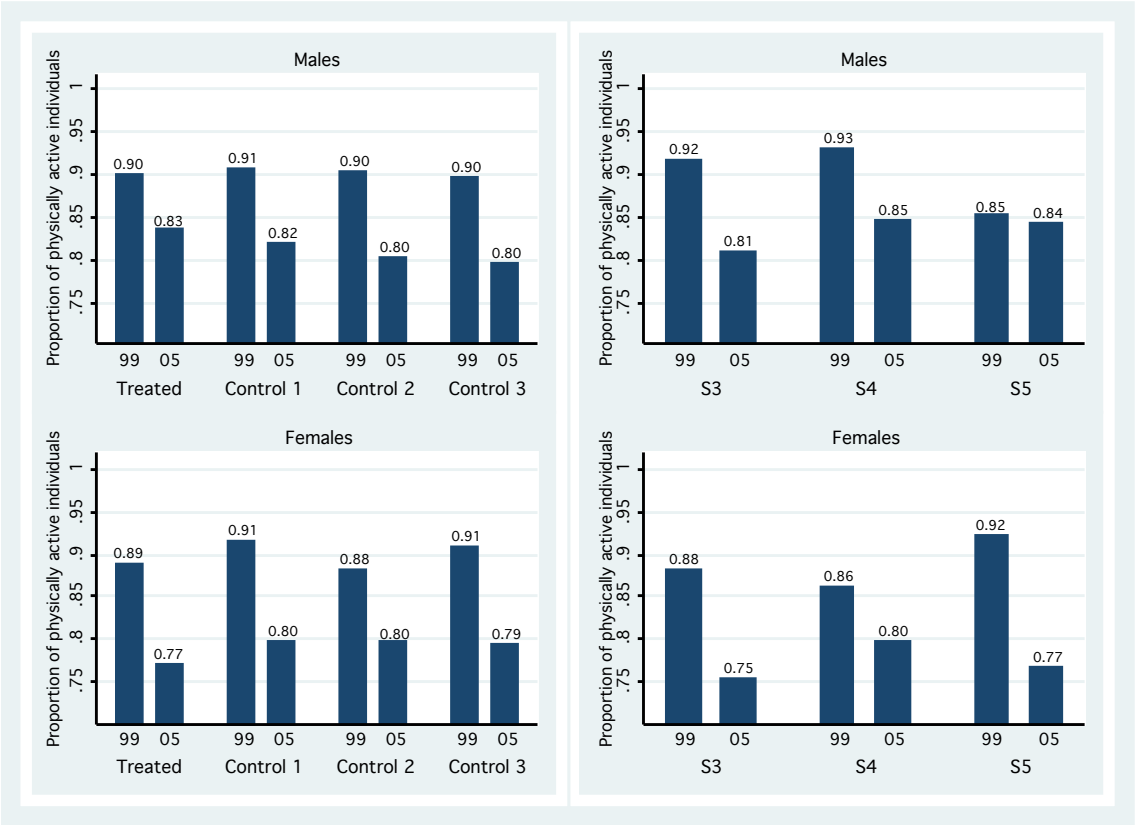
Source: PSID

downward trend in all groups for both genders. In particular for the treated groups, the proportion of physically active individuals goes down by 7 percentage points for males, and by 12 percentage points for females. This simple Before-After estimator tells us that HED policies have had a negative impact on the outcome of interest. However, these estimates are obviously biased given that the average of the outcome variable in the three control groups also has a downward trend.

Exploring gender differences, we can see that females in the Treatment-Experimental group (Treated) present a larger drop in the proportion of physically active individuals than that observed for males in the same group. This suggests the need to take gender differences into account when estimating the effect of HED policies.

As we discussed above, the implementation and modification of HED policies between 1999 and 2005 were not homogeneous across states. Therefore, we can expect differences in the temporal evolution of the outcome of interest for treated individuals across the three groups of experimental states. The two graphs in the right panel in Figure 2.2 show the proportion of physically active treated individuals, by gender and by group of experimental states. In the first graph we see that in states belonging to group $S5$, the states that introduced major HED changes, the downward trend in the proportion of physically active males is substantially smaller than the corresponding downward trend in groups $S3$ and $S4$, the groups of states that introduced moderate HED changes. Moreover, the reduction in the proportion of physically active males in the group $S5$ is lower than the fall in all three control groups. This relatively moderate downward trend for treated males in $S5$

Figure 2.2: Proportion of physically active individuals by treated/control groups (left panel), and treated individuals by treatment groups (right panel), and by gender, in 1999 and 2005.



Notes: **Treated:** individuals with elementary school-age children in experimental states. Control 1: individuals with elementary school-age children in non-experimental states. **Control 2:** individuals without elementary school-age children in experimental states. **Control 3:** individuals without elementary school-age children in non-experimental states. The type of policies corresponding to the groups of states Sk are as follows. **S3:** Moderate changes A; **S4:** Moderate changes B; **S5:** Major changes. Source: PSID.

experimental states suggests a positive effect of HED policies on the outcome variable, although it does not seem to be the case for females.

2.3.3 DDD estimation in a simple linear model

Table 2.4 presents the DDD estimate of the effect of changes in HED policy on the behavior of fathers for the group of states that introduced major HED changes, $S5$.[21]

Panel A compares the change in the proportion of physically active males with elementary school-age children residing in $S5$ states to the change for the group of fathers with elementary school-age children in non-experimental $S1$ or $S2$ states. Each cell in the first two columns contains the proportion of physically active individuals for the corresponding

[21] In Table 2.14 in the Appendix we report results for a similar exercise on mothers.

group, before and after the HED reform, along with the standard errors and the number of observations. The Before-After estimate (Δ_E^T) of the effect is shown in the third column. There is a non-significant decrease in the proportion of physically active fathers with elementary school-age children in experimental states, and a significant fall in the proportion of physically active fathers with children of the same age in non-experimental states. The diff-in-diff estimator ($\Delta_E^T - \Delta_{NE}^T$), reported at the bottom of Panel A, is positive but non-significant. However, the DD estimate could be downwardly biased because, as we showed with the data from the BRFSS, the policy changes occurred in those states where health outcomes and health behaviors were deteriorating more rapidly. Hence, to control for potential state-specific trends, we additionally look at the evolution of outcomes for a control group within each state.

Table 2.4: DDD estimator for males in $S5$.

	Before HED change	**After HED change**	**Time difference**	
A. Treatment individuals: with elementary school-age children				
Experimental states	0.851	0.843	-0.009	Δ_E^T
	(0.036)	(0.035)	(0.050)	
	[101]	[108]		
Non-experimental states	0.908	0.821	-0.087***	Δ_{NE}^T
	(0.012)	(0.016)	(0.020)	
	[606]	[563]		
Difference in difference			**0.078**	
			(0.054)	
B. Control Individuals: without elementary school-age children				
Experimental states	0.918	0.766	-0.152***	Δ_E^C
	(0.023)	(0.030)	(0.038)	
	[147]	[201]		
Non-experimental states	0.896	0.803	-0.093***	Δ_{NE}^C
	(0.011)	(0.012)	(0.016)	
	[834]	[1,162]		
Difference in difference			**-0.059**	
			(0.041)	
C. Non-parametric DDD estimator				
$\mathbf{DDD = (\Delta_E^T - \Delta_{NE}^T) - (\Delta_E^C - \Delta_{NE}^C)}$			**0.138****	
			(0.068)	

Notes: Cells contain proportion of physically active individuals for the group identified. Standard errors are given in parentheses, and sample sizes in brackets. The non-experimental states are groups of states S1 and S2. Significance levels: * = 10%; ** = 5%; *** = 1%.

In Panel B of Table 2.4 we perform the same exercise for the groups of fathers with children above and below elementary school age. We find a larger fall in the proportion of physically active individuals in the experimental states, relative to the other states, as was expected

according to the pre-treatment trends observed in the BRFSS data.

In panel C we compute the difference between the two DD estimators in panel A and B. This non-parametric DDD estimator indicates that there is a 13.8 percentage points significant increase in the relative proportion of physically active fathers of elementary school-age children, compared to the change in the relative proportion of physically active fathers with no elementary school-age children. This statistically significant DDD estimate provides some evidence on the existence of spillovers of HED on the physical activity of fathers.

In the following subsections we discuss how the DDD design can be expressed in a regression framework that will allow us to control for observable differences between individuals in the treated and control groups, as well as explicitly modeling the discrete support of the outcome variable.

2.3.4 Allowing for covariates and gender differences

In Table 2.5 we report average values and standard errors of the outcome variable, and other demographic and socioeconomic characteristics for treated and control individuals in 1999 and 2005.

For each group, we find evidence of statistically significant differences in some observable characteristics between 1999 and 2005. These differences may produce changes in the observed proportion of physically active individuals between 1999 and 2005 that are not a consequence of changes in HED programs. To avoid a biased estimation of the effect of interest, we use a regression framework that allows us to control for temporal differences in observable characteristics.

Given the existence of different time trends on the frequency of light physical activity between females and males observed in Figure 2.2, in the model that we estimate we include interactions of all the parameters related to the identification of the HED effect with a gender dummy.

The outcome equation with interactions by gender and with covariates has the following form

$$\begin{aligned} y_{it} = {} & \beta_0 + \beta_1\tau_t + \beta_2 elem_i + \sum_{k=2}^{5}\beta_{3,k}Sk_i + \beta_4 female_i + \beta_5(\tau_t \times female_i) + \beta_6(elem_i \times female_i) + \\ & \sum_{k=2}^{5}\beta_{7,k}(Sk_i \times female_i) + \beta_8(elem_i \times \tau_t) + \beta_9(elem_i \times \tau_t \times female_i) + \sum_{k=2}^{5}\beta_{10,k}(Sk_i \times \tau_t) + \\ & \sum_{k=2}^{5}\beta_{11,k}(Sk_i \times \tau_t \times female_i) + \sum_{k=2}^{5}\beta_{12,k}(Sk_i \times elem_i) + \sum_{k=2}^{5}\beta_{13,k}(Sk_i \times elem_i \times female_i) + \\ & \sum_{k=3}^{5}\beta_{14,k}(Sk_i \times elem_i \times \tau_t) + \sum_{k=3}^{5}\beta_{15,k}(Sk_i \times elem_i \times \tau_t \times female_i) + \beta_{16}X_{it} + u_{it}, \end{aligned} \tag{2.6}$$

Table 2.5: Descriptive statistics: All individuals in the sample.

	Treated individuals			Control individuals		
	1999	2005	Difference	1999	2005	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Frequency of light physical	4.31	3.82	-0.49***	4.37	3.75	-0.62***
activity (times per week)	(3.09)	(3.26)		(3.09)	(3.23)	
Body Mass Index	27.11	27.59	0.48	26.53	27.74	1.21***
	(5.77)	(5.67)		(5.30)	(5.96)	
Proportion with Health condition	0.11	0.14	0.03*	0.14	0.17	0.03***
that limits daily activity	(0.32)	(0.35)		(0.34)	(0.37)	
Proportion of Female	0.56	0.57	0.01	0.55	0.57	0.02*
	(0.50)	(0.50)		(0.50)	(0.50)	
Age	36.13	36.15	0.02**	37.17	39.46	2.29***
	(6.34)	(6.84)		(8.25)	(9.92)	
Years of Education completed	13.01	13.22	0.21	12.79	13.05	0.26***
	(2.38)	(2.25)		(2.74)	(2.49)	
Num. of Children	2.62	2.58	-0.04	2.34	2.31	-0.03
	(1.33)	(1.26)		(1.25)	(1.20)	
Num. of Children	1.26	1.28	0.03	0.45	0.31	-0.14***
in elementary school	(0.50)	(0.52)		(0.72)	(0.60)	
Proportion of White	0.53	0.50	-0.03	0.56	0.54	-0.02*
	(0.50)	(0.50)		(0.50)	(0.50)	
Proportion of Married	0.77	0.75	-0.02	0.78	0.76	-0.02**
	(0.42)	(0.43)		(0.42)	(0.43)	
Proportion of Unemployed	0.04	0.04	0.00	0.03	0.04	0.01**
	(0.21)	(0.19)		(0.17)	(0.19)	
Proportion of Retired	0.00	0.00	0.00	0.00	0.01	0.01***
	(0.04)	(0.05)		(0.07)	(0.10)	
Proportion of Disabled	0.01	0.02	0.00	0.02	0.03	0.01***
	(0.12)	(0.14)		(0.14)	(0.17)	
Proportion of Full time	0.74	0.70	-0.03	0.77	0.73	-0.03***
workers	(0.44)	(0.46)		(0.42)	(0.44)	
Labor income	13,621	17,723	4,102***	14,878	19,765	4,887***
per capita	(16,916)	(28,353)		(15,911)	(29,240)	
Total income	16,392	27,805	11,413***	17,721	23,605	5,884***
per capita	(20,049)	(214,424)		(21,151)	(33,232)	
Sample size	679	661		4,061	5,262	

Notes: Standard errors reported in parentheses below the corresponding average. Stars in columns (3) and (6) show statistical significance of differences in proportion or distribution of the referred variable, between years 1999 and 2005. We perform tests of difference in proportion for the dummy variables White, Health condition that limits daily activity, Married, Unemployment, Retired, Disabled, and Full-time workers. We perform tests of differences in distribution for the categorical variables Frequency of light physical activity, Age, Education, Number of Children, and Number of Children in elementary school, and for the continuous variables Body Mass Index, Labor income, and Total income. Significance levels: * = 10%; ** = 5%; *** = 1%.

where $i = 1...N$ indexes individuals, $t = 0, 1$ indexes time (0=before policy, 1999; 1=after policy, 2005), and $k = 1, ..., 5$ indexes state groups.[22]
The DDD estimates in this model are the estimates of $\beta_{14,k}$ for males, and $\beta_{14,k} + \beta_{15,k}$ for females. If the parameter $\beta_{15,k}$ is significantly different from zero, there is evidence of a different impact of HED policies between fathers and mothers. X_{it} is a set of observable individual characteristics including age, race, gender, health conditions that limits daily activity, body mass index, marital status, number of children, children of high-school-age, education level, employment status, full-time/part-time employment, total family income, and state of residence.

2.3.5 Empirical implementation: DDD in a non-linear model

To simplify notation, in this section we use the specification of the outcome equation in (2.1), which does not include state classification, covariates, and gender interactions.
Considering that the outcome variable is binary, the expectation of the outcome equation measures the probability of doing light physical activity any positive number of times per week, and has the following form

$$\begin{aligned}\mathbb{E}[y_{it}|elem_i, S_i, \tau_t] = f\Big[&\beta_0 + \beta_1\tau_t + \beta_2 elem_i + \beta_3 S_i \\ &+ \beta_4(elem_i \times \tau_t) + \beta_5(S_i \times \tau_t) + \beta_6(elem_i \times S_i) \\ &+ \beta_7(\tau_t \times elem_i \times S_i)\Big],\end{aligned} \tag{2.7}$$

where f is the cumulative distribution function of idiosyncratic shocks (u_{it}).
As remarked in Blundell and Dias (2009), applying DD and DDD methods imposes additive separability of the error term conditional on the observables, an assumption that does not hold when the outcome of interest is a dummy variable. To overcome this limitation, we follow Blundell, Dias, Meghir, and Reenen (2004) by imposing the identifying assumption in equation (2.3) over the index, rather than over the probability itself. Assuming that the inverse probability function, f^{-1}, is known, the DDD estimator of the ATT is

$$\begin{aligned}\widehat{ATT} = \bar{y}_1^{1,1} - f\Big\{&f^{-1}(\bar{y}_0^{1,1}) + [f^{-1}(\bar{y}_1^{1,0}) - f^{-1}(\bar{y}_0^{1,0})] \\ &+ [f^{-1}(\bar{y}_1^{0,1}) - f^{-1}(\bar{y}_0^{0,1})] - [f^{-1}(\bar{y}_1^{0,0}) - f^{-1}(\bar{y}_0^{0,0})]\Big\},\end{aligned} \tag{2.8}$$

where $\bar{y}_t^{elem,S}$ is the average of the estimated outcome over individuals in group $elem$, residing in states S, at time t.[23]
Assuming that the idiosyncratic shocks have a normal distribution, f is the normal cumu-

[22] $S1$ is the group of reference.
[23] In Section 2.7.1 in the Appendix we show how we obtain the expression for the DDD estimator.

lative distribution function. We estimate the parameters of interest by maximum likelihood and compute robust standard errors clustered at the family level. A report of the estimated coefficients can be found in Table 2.15 in the Appendix.
With the estimated parameters we compute the Indirect Average Treatment effects on the Treated (IATT), using equation (2.8), including the state classification (discussed in Section 2.3.1), gender interactions, and covariates (both presented in Section 2.3.4).

2.4 IATT estimates

Table 2.6 shows the IATT for the three types of treatment, by gender. The "OLS" column presents the IATT estimates obtained using a linear probability model. The "Probit" column presents the IATT estimates obtained using equation (2.8) and assuming a normal distribution for idiosyncratic shocks.

Table 2.6: IATT by type of treatment, and by gender.

Group of experimental states	**Male**			**Female**		
	OLS	**Probit**	**# obs**	**OLS**	**Probit**	**# obs**
S3: Moderate changes A	-0.009	-0.039	3,628	-0.042	-0.028	4,595
	(0.071)	(0.080)		(0.065)	(0.085)	
S4: Moderate changes B	-0.035	-0.055	3,678	0.023	0.034	4,602
	(0.064)	(0.065)		(0.059)	(0.081)	
S5: Major changes	**0.122***	**0.199***	3,722	-0.056	-0.088	4,686
	(0.068)	(0.107)		(0.054)	(0.061)	

Notes: Robust standard errors reported in parenthesis clustered at the family level. Robust standard errors computed by bootstrap using 1000 replications in Probit specification."OLS" columns present the IATT estimates obtained using a linear probability model. The "Probit" columns present the IATT estimates obtained using equation (2.8) and and probit specification. The regressions include the following covariates: age, race, gender, health status, marital status, number of children, children of high school-age, education level, employment status, full-time/part-time employment, total family income level, and state of residence. Significance levels: * = 10%; ** = 5%; *** = 1%.

We find evidence of a positive effect of HED education at the elementary school level on the probability of parents engaging in light physical activity. A noteworthy change in the HED program ($S5$ group of states) raises the probability of fathers doing physical activity. Looking at the results of the probit column we can see that the probability of fathers affected by this policy doing physical activity is 19.9 percentage points higher than that of fathers not affected by the policy. The positive and statistically significant effect on fathers is also obtained by using a linear probability model. The effect on the probability that mothers engage in light physical activity is never statistically significant, but the signs are the opposite to those found for fathers.
The estimated effects are not statistically significant for males and females residing in the

group of states $S3$ and $S4$. These results suggest that moderate changes in HED programs do not produce indirect effects.

The interpretation of the estimated IATT can be clarified by looking at the averages of the estimated outcomes in Table 2.7. Let's consider the results for treated fathers residing in the group of states S_5. On average, the estimated percentage of physically active fathers in the pre-treatment period, 1999, is 84.9%. In 2005, after major changes in HED programs, we estimate that 84.3% of fathers were engaged in light physical activity. Nevertheless, if HED programs had not been subject to profound changes in this group of states, we estimate that only 64.4% of fathers of elementary school-age children would have engaged in light physical activity in 2005. In other words, due to the major changes in HED programs, the percentage of physically active fathers fell from 84.8% to 84.3%, instead of falling to 64.4% had HED not been modified. The effect of major reforms on HED was to soften the declining trend in the proportion of physically active fathers.

Table 2.7: IATT and averages of the estimated outcomes for treated individuals, by type of treatment and gender.

Group of experimental states		**IATT**	**Estimated Average Outcomes for Treated Individuals**		
			Post-treatment period with treatment	Post-treatment period without treatment	Pre-treatment period without treatment
S3: Moderate change A	Male	-0.039	0.810	0.848	0.919
	Female	-0.028	0.759	0.787	0.884
S4: Moderate change B	Male	-0.055	0.848	0.903	0.928
	Female	0.034	0.796	0.762	0.863
S5: Major change	Male	**0.199**	**0.843**	**0.644**	**0.849**
	Female	-0.088	0.766	0.854	0.924

Notes: IATT and estimated outcomes obtained using equation (2.8) and a probit specification. Each cell contains the estimated proportion of physically active individuals. Pre-treatment period is 1999, and post-treatment period is 2005. The IATT is obtained as the estimated average outcome in the post-treatment period under treatment minus the missing counterfactual, that is, the estimated average outcome in the post-treatment period without treatment.

We conclude that there are positive spillovers of introducing major changes in existing HED programs on the probability of fathers engaging in light physical activity, while for mothers we do not find a statistically significant effect of these reforms.

2.4.1 Plausible explanations for our results

We can think of two channels to explain our results. When children start receiving HED at school their parents are confronted with two new sets of factors that might potentially affect their health-related behaviors. First, parents may optimally react to HED in schools by complementing this education with the incorporation of healthy lifestyles into their own

daily activities. We refer to this potential channel as "role modeling". On the other hand, there is the effect of the arrival of new information that the child receives at the school. In particular, parents are confronted with knowledge that the child brings to the household from the health education curricula given at the school, and they may adjust their health behaviors in response to it. We refer to this potential channel as "information sharing". In what follows we provide evidence of the existence of both channels.

2.4.1.1 Role models

Parents may do more physical exercise in response to the knowledge children acquire via HED, not because they were not already aware of the benefits of exercising but because they want to complement the instruction received by the child so as to form the desired healthy lifestyle in the child.

The estimates from the model interacting the policies with a dummy variable for gender provide some insights on the operation of the "role model" channel. Parents usually spend more time with their children doing gendered activities. Figure 2.3 in the Appendix shows some evidence on this respect with data from the American Time Use Survey (ATUS). Women spend roughly twice as much time in childcare as do men, a pattern which holds true for all subgroups and for almost all types of childcare, except for "Recreational" childcare. This type of childcare activity includes playing games with children, playing outdoors with children, attending a child's sporting event or dance recital, going to the zoo with children, taking walks with children, etc. In the case of "Recreational" childcare, mothers allocate relatively less of their time with children than do fathers. Thus, this is evidence that fathers are more likely to do stereotypically male activities with their children, among them physical activity. Accordingly, the impact of HED reforms on physical activity is more likely to appear for fathers rather than for mothers.

2.4.1.2 Information sharing between children and parents

Individuals with a lower stock of information are expected to be more affected by HED changes. We explore the existence of the information-sharing channel by analyzing the differential impact of HED reforms on individuals with low and high education levels and with low and high income levels. Since lower levels of education and socioeconomic status are related to less knowledge about health (Kenkel, 1991; Tinsley, 2003), we expect to obtain a greater effect of HED reforms on individuals with lower levels of education and income.

Exploiting the non-linearity of the model specified we estimate the IATT evaluated at particular values of the covariates of interest. We report the results in Table 2.8 for treated fathers residing in states that belong to group $S5$. According to these results, the policy has a higher effect on lower educated males relative to higher educated males, on non-white

males relative to white males, and on males that have a lower income than those having a higher income. The policy raises the probability of lower educated males being physically active relative to higher educated males in 4.9 percentage points, while the increment is 4.2 percentage points for non-white males relative to white males.

Table 2.8: Differences in IATT estimates evaluated at particular values of the covariates, for males in $S5$.

Income	**IATT**	**Education**	**IATT**	**Race**	**IATT**
Low	0.198*	**Low**	0.211*	**No White**	0.207*
(20th percentile)	(0.108)		(0.114)		(0.116)
High	0.193*	**High**	0.162*	**White**	0.165*
(80th percentile)	(0.106)		(0.096)		(0.099)
Difference	**0.005***		**0.049****		**0.042****
	(-0.003)		(0.021)		(0.020)

Notes: The IATT is obtained using equation (2.8) and a probit specification. Robust standard errors in parentheses clustered at the family level, and computed by bootstrap using 1000 replications. We find no differences for males in health status (measured using the existence of a health condition that limits the daily activity and the body mass index), family size, labor force participation, full-time/part-time employment, and marital status. There are no differences for females in all the dimensions analyzed. Significance levels: * = 10%; ** = 5%; *** = 1%.

2.5 Robustness

2.5.1 Validity of the identifying assumption

Our identifying assumption requires that, in the absence of HED reforms, state specific trends of the proportion of parents physically active in the treatment group (those with elementary school-age children) is the same as that of parents in the control group (those with children bellow or above elementary school-age). This assumption will be violated if there is a shock contemporaneous to HED reforms that systematically affects the relative outcome of parents of elementary school-age children.

In order to check the robustness of our identifying assumption we perform two tests. First, we estimate the effect of the HED reform on the labor force participation of parents. Labor participation is a decision that should be not affected by the policy under analysis. If HED changes are not the only shock that affects the relative outcomes of treated versus control individuals between 1999 and 2005, we may observe that the labor force participation among individuals in the treatment group changes relative to that of individuals in the control group. Results for the group of states $S5$ are presented in Table 2.9. The two estimates tell us that between 1999 and 2005 there were no significant changes in the relative decision of participating in the labor market of treated and control parents. This constitutes evidence that there was no other labor market related shock systematically affecting the relative outcomes of treated and control parents between 1999 and 2005.

Second, we estimate the effect of HED reforms in elementary school on individuals that are

not likely to be affected by such reforms: the group of adults without children. If there had been no other shock contemporaneous to the HED reforms, the outcome of adults without children relative to a control group should remain unchanged between 1999 and 2005. To perform this test, we keep only individuals without children. Individuals in a given state are then assigned to one of two groups, the "placebo" treatment and the "placebo" control group. The classification is done in such a way that the "placebo" treatment group resembles, at least in observable characteristics, the true treatment group of adults with children in elementary school. This classification is necessary to have a within state control group (the third dimension) in the sample of adults without children, required to implement a DDD estimator. If our classification is correct, we should find no effect of HED reforms on the "placebo" treated individuals. Details on how we select the "placebo" treated individuals can be found in the Appendix in Section 2.7.2. We present the results using the non-parametric DDD estimator computed in Section 2.3.3 for the group of states S_5.[24] The effect of HED on the "placebo" treated males is 0.062 (standard error: 0.097), and the effect on the "placebo" females is -0.047 (standard error: 0.097). Reassuringly, the estimated coefficients are smaller than those in the baseline model and not significantly different from zero.[25]

2.5.2 Sensitivity analysis

Parents in the treatment group, that is parents of elementary school-age children, may also have other children below and/or above elementary school age. In order to have individuals in the control group comparable to those in the treatment group in the same state, we consider that the appropriate group of control individuals should include parents of children of ages below and above elementary school age. Nevertheless, to determine whether our results are sensitive to this definition, we perform two tests. First, we use parents with at lest one child below elementary-school-age as non-treatment individuals. Second, we use parents with at least one child above elementary-school-age as non-treatment individuals. Estimates of the effects of HED reforms on the group of states $S5$ are very similar to those obtained with the baseline model, as can be seen in Table 2.9.[26]

2.6 Conclusion

We find evidence for positive spillovers of HED imparted in elementary schools on the probability of parents engaging in light physical activity. However, our results suggest

[24] We cannot obtain the estimates of the probit or OLS models because the sample size of adults without children is not large enough.

[25] We also perform the placebo test for the group of states $S3$ and $S4$ and results and conclusions are similar to those obtained with the group of states $S5$.

[26] We perform both tests for the group of states $S3$ and $S4$ and results and conclusions are similar to those obtained with the group of states $S5$.

Table 2.9: Robustness check: Effect of HED reforms on labor market participation (Panel A), and Sensitivity of control groups (Panel B and C), for group of states $S5$.

	Male			**Female**		
	OLS	**Probit**	**# obs**	**OLS**	**Probit**	**# obs**
	PANEL A: *Effect of HED reforms on* ***labor market participation***					
IATT	0.029	0.107	3,722	0.029	0.017	4,686
	(0.054)	(0.098)		(0.065)	(0.066)	
	PANEL B: *Using as* ***control*** *individuals only parents of children* ***below*** *elementary-school-age*					
IATT	0.098	0.220	2,360	-0.012	-0.010	2,930
	(0.073)	(0.144)		(0.070)	(0.093)	
	PANEL C: *Using as* ***control*** *individuals only parents of children* ***above*** *elementary-school-age*					
IATT	0.135*	0.188	2,894	-0.083	-0.117*	3,735
	(0.081)	(0.113)		(0.061)	(0.062)	

Notes: The IATT in columns "OLS" are obtained using a linear probability model. The IATT in columns "Probit" are obtained using equation (2.8) and a probit specification. Robust standard errors reported in parenthesis clustered at the family level. Robust standard errors computed by bootstrap using 1000 replications in Probit specification. The regressions in Panel A include the following covariates: age, race, gender, health status, marital status, number of children, children of high school-age, education level, total family income level, and state of residence. The regressions in Panel B and Panel C include the previous covariates and employment status, and full-time/part-time employment. Significance levels: * = 10%; ** = 5%; *** = 1%.

that fathers and not mothers are those affected by the HED reforms. We also analyze the differential impact of HED reforms on fathers and mothers as a way to explore the nature of the channels driving the spillovers.

We argue that a "role model" channel can explain the differential impact on fathers and mothers. The idea is based on the different role models that mothers and fathers play for their children. Parents usually spend more time with their children doing gendered activities. Since physical activity can be included in the group of typically male-activities, the effect of promoting the advantages of doing physical activity is more likely to appear for fathers rather than for mothers. We also explore the existence of a second channel driving our results -the "information sharing" channel- by analyzing the differential impact of HED reforms on individuals with lower and higher education levels, and obtain the expected greater effect on less educated individuals and individuals with lower socioeconomic status. Our results also highlight the importance of clearly distinguishing the existence of several dimensions in the implementation of a policy. In our case, it is important for policy evaluation to consider the two dimensions in HED reforms, changes in topics and enforcements, as well as the distinction between "Moderate changes" and "Major changes" in HED requirements. Our main result shows spillovers only in states that carried out profound reforms in their HED programs.

Spillovers of HED on parents' lifestyles indicate that the interaction between children and parents plays a role in the formation of healthy lifestyles within the household. Therefore,

taking these spillovers into account is important in the cost-benefit analysis of health education in schools. In addition, the conclusion that implementing minor reforms in existing HED programs is not enough to obtain spillovers at the family level helps to properly design policy interventions aimed at increasing the adoption of healthy lifestyles.

2.7 Appendix

Table 2.10: HED topics and enforcements. Full list.

Topics List
1) *Alcohol- or Other Drug-Use Prevention*
2) *Emotional and Mental Health*
3) *Nutrition and Dietary Behavior*
4) *Physical Activity and Fitness*
5) *Tobacco-Use Prevention*
6) Human immunodeficiency virus (HIV) prevention
7) Accident or injury prevention
8) Sexually transmitted disease (STD) prevention
9) Pregnancy prevention
10) Suicide prevention
11) Violence prevention, for example bullying, fighting, or homicide

Enforcements List
1) *State requires districts or schools to follow national or state health education standards or guidelines*
2) *State requires students in elementary school to be tested on health topics*
3) *State requires each school to have a HED coordinator*
4) State uses staff development for HED teachers to improve compliance with HED standards or guidelines
5) State uses written reports from districts or schools to document compliance with HED standards or guidelines
6) State provides a list of one or more recommended elementary school HED curricula
7) State provides a chart describing the scope and sequence of instruction for elementary school HED
8) State provides lesson plans or learning activities for elementary school HED
9) State provides plans for how to assess or evaluate students in elementary school HED
10) State adopts a policy stating that newly hired staff who teach HED at the elementary school level will have undergraduate or graduate training in HED
11) State offers certification, licensure, or endorsement to teach HED
12) State adopts a policy stating that teachers will earn continuing education credits on HED topics to maintain state certification, licensure, or endorsement to teach HED

Notes: The topics and enforcements considered for the analysis are in italics.

Table 2.11: HED programs: health topics required, by state and year.

	1999					2005				
State	**topic 1**	**topic 2**	**topic 3**	**topic 4**	**topic 5**	**topic 1**	**topic 2**	**topic 3**	**topic 4**	**topic 5**
Alabama	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Alaska	no	no	no	no	no	no	no	no	no	no
Arizona	no	no	no	no	no	no	no	no	no	no
Arkansas	no	no	no	no	no	yes	no	yes	yes	yes
California	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Colorado	no	no	no	no	no	no	no	no	no	no
Connecticut	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Delaware	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
District of Columbia	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Florida	no	no	no	no	no	yes	yes	yes	yes	yes
Georgia	yes	yes	yes	no	yes	yes	yes	yes	yes	yes
Hawaii	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Idaho	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Illi0is	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Indiana	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Iowa	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Kansas	.	.	.	.	.	.	.	.	.	.
Kentucky	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Louisiana	yes	no	no	no	yes	yes	yes	yes	yes	yes
Maine	yes	yes	yes	no	yes	yes	yes	yes	yes	yes
Maryland	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Massachusetts	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Michigan	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Minnesota	.	.	.	.	.	.	.	.	.	.
Mississippi	no	no	no	no	no	no	no	no	no	no
Missouri	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Montana	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Nebraska	yes	no	no	no	yes	yes	no	yes	yes	yes
Nevada	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
New Hampshire	.	.	.	.	.	yes	yes	yes	yes	yes
New Jersey	yes	no	no	yes	yes	yes	no	yes	yes	yes
New Mexico	no	no	no	no	no	yes	yes	yes	yes	yes
New York	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
0rth Carolina	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
0rth Dakota	yes	no	no	no	yes	yes	no	no	yes	yes
Ohio	yes	no	yes	yes	yes	yes	no	yes	yes	yes
Oklahoma	no	no	no	no	no	no	no	no	no	no
Oregon	yes	no	no	no	yes	yes	no	no	no	yes
Pennsylvania	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Rhode Island	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
South Carolina	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
South Dakota	no	no	no	no	no	no	no	no	no	no
Tennessee	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Texas	.	.	.	.	.	yes	yes	yes	yes	yes
Utah	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Vermont	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Virginia	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Washington	yes	no	yes	yes	yes	yes	yes	yes	yes	yes
West Virginia	yes	no	yes	yes	yes	yes	no	yes	yes	yes
Wisconsin	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wyoming	no	no	no	no	no	no	yes	no	no	no

Source: NASBE Database and School Health Policies and Programs Study (SHPPS).

Notes: The data contained in this table was constructed cross-checking the information from both sources, and in most of the cases survey information from SHPPS coincides with the legal information summarized in NASBE. In those cases in which there is no coincidence, we rely on NASBE information only. In few cases NASBE does not provide complete information -i.e., cases in which the regulations contained in NASBE are not informative about the characteristics of the policy the state implements-, then we rely on SHPPS. Missing values indicate that the information cannot be recovered from any of the two sources.

Topic 1:Alcohol or other drug-use prevention; **Topic 2**: Emotional and mental health; **Topic 3**: Nutrition and dietary behavior; **Topic 4**: Physical activity and fitness; **Topic 5**: Tobacco-Use prevention.

Table 2.12: HED programs: enforcements required, by state and year.

	1999			2005		
State	**enf 1**	**enf 2**	**enf 3**	**enf 1**	**enf 2**	**enf 3**
Alabama	yes	no	no	yes	no	yes
Alaska	no	no	no	no	no	no
Arizona	yes	no	no	yes	no	no
Arkansas	yes	no	no	yes	no	no
California	no	no	no	no	no	no
Colorado	no	no	no	no	no	no
Connecticut	no	no	no	no	no	no
Delaware	yes	no	yes	yes	no	yes
District of Columbia	.	.	.	.	.	.
Florida	yes	no	no	yes	no	no
Georgia	yes	no	no	yes	no	no
Hawaii	yes	no	no	yes	no	no
Idaho	no	no	no	no	no	no
Illinois	yes	no	no	yes	no	no
Indiana	yes	no	no	yes	no	no
Iowa	no	no	no	no	no	no
Kansas	.	.	.	.	.	.
Kentucky	no	yes	no	yes	yes	no
Louisiana	yes	no	no	yes	no	no
Maine	yes	yes	no	yes	yes	no
Maryland	yes	no	no	yes	no	no
Massachusetts	yes	no	no	yes	no	no
Michigan	yes	no	no	yes	no	no
Minnesota	.	.	.	.	.	.
Mississippi	no	no	no	no	no	no
Missouri	yes	yes	no	yes	yes	no
Montana	yes	no	no	yes	no	no
Nebraska	no	no	no	no	no	no
Nevada	yes	no	no	yes	no	no
New Hampshire	.	.	.	yes	.	.
New Jersey	.	.	.	yes	.	.
New Mexico	no	.	.	yes	.	.
New York	no	.	.	no	.	.
north Carolina	yes	no	no	yes	no	no
north Dakota	no	no	no	no	no	no
Ohio	.	.	.	.	.	.
Oklahoma	no	no	no	no	no	no
Oregon	no	no	no	yes	no	no
Pennsylvania	yes	.	.	yes	.	.
Rhode Island	yes	yes	no	yes	yes	yes
South Carolina	yes	no	no	yes	yes	no
South Dakota	no	no	no	no	no	no
Tennessee	yes	no	no	yes	no	no
Texas	no	no	no	yes	no	no
Utah	yes	no	no	yes	yes	no
Vermont	yes	no	no	yes	yes	no
Virginia	yes	no	no	yes	no	no
Washington	yes	yes	no	yes	yes	no
West Virginia	yes	no	no	yes	no	no
Wisconsin	no	no	no	no	no	no
Wyoming	no	no	no	yes	no	no

Source: NASBE Database and School Health Policies and Programs Study (SHPPS).
Notes: The data contained in this table was constructed cross-checking the information from both sources, and in most of the cases survey information from SHPPS coincides with the legal information summarized in NASBE. In those cases in which there is no coincidence, we rely on NASBE information only. In few cases NASBE does not provide complete information -i.e., cases in which the regulations contained in NASBE are not informative about the characteristics of the policy the state implements-, then we rely on SHPPS. Missing values indicate that the information cannot be recovered from any of the two sources.
Enforcement 1: State requires districts or schools to follow national or state health education standards or guidelines.
Enforcement 2: State requires students in elementary school to be tested on health topics.
Enforcement 3: State requires each school to have a HED coordinator.

Table 2.13: States classified by groups Sk.

NON-EXPERIMENTAL		**EXPERIMENTAL**	
State	# of obs.	State	# of obs.
S1		**S3**	
Alaska	14	Alabama	138
Colorado	246	Kentucky	169
Mississippi	521	Oregon	208
Oklahoma	62	Rhode Island	10
South Dakota	59	South Carolina	564
		Utah	100
S2		Vermont	7
Arizona	178		
California	1,218	**S4**	
Connecticut	79	Georgia	414
Delaware	14	Louisiana	209
Hawaii	4	Maine	30
Idaho	25	Nebraska	106
Illinois	397	New Jersey	336
Indiana	363	North Dakota	16
Iowa	264	Washington	218
Kansas	81		
Maryland	450	**S5**	
Massachusetts	258	Arkansas	278
Michigan	613	Florida	450
Missouri	340	New Mexico	16
Montana	13	Texas	691
Nevada	72	Wyoming	18
New York	493		
North Carolina	605		
Ohio	505		
Pennsylvania	476		
Tennessee	238		
Virginia	373		
West Virginia	24		
Wisconsin	183		

Notes: We do not include the District of Columbia, Minnesota, and New Hampshire since the information regarding HED policies for these states is not precise in terms of when HED was implemented, making impossible their classification.

Table 2.14: DDD estimator for females in $S5$.

	Before HED change	**After HED change**	**Time difference**	
A. Treatment individuals: with children in elementary school				
Experimental states	0.922	0.767	-0.155***	Δ_E^T
	(0.024)	(0.035)	(0.042)	
	[128]	[146]		
Non-experimental states	0.914	0.798	-0.116***	Δ_{NE}^T
	(0.010)	(0.015)	(0.018)	
	[758]	[753]		
Difference in difference			**-0.039**	
			(0.046)	
B. Control Individuals: without children in elementary school				
Experimental states	0.874	0.785	-0.089**	Δ_E^C
	(0.025)	(0.025)	(0.035)	
	[183]	[270]		
Non-experimental states	0.910	0.796	-0.114***	Δ_{NE}^C
	(0.009)	(0.011)	(0.014)	
	[984]	[1,464]		
Difference in difference			**0.025**	
			(0.038)	
C. Non-parametric DDD estimator				
$\mathbf{DDD = (\Delta_E^T - \Delta_{NE}^T) - (\Delta_E^C - \Delta_{NE}^C)}$			**-0.063**	
			(0.060)	

Notes: Cells contain proportion of physically active individuals for the group identified. Standard errors are given in parentheses, and sample sizes in brackets. The non-experimental states are groups of states S1 and S2. Significance levels: * = 10%; ** = 5%; *** = 1%.
The upper part of the Table shows important falls in temporal trends of the proportions of physically active mothers of elementary school-age children residing in both, experimental and non-experimental states. As a consequence, the difference-in-difference estimator is not statistically significant. We can observe a similar pattern for mothers of children below and above elementary school age. Finally, the triple difference estimator does not provide evidence of an effect of HED on the proportion of physically active mothers.

Table 2.15: Probit Model: Probability of doing light physical activity at least once a week.

Number of obs= 10,663 — Wald chi2(83) = 795.28
Prob > chi2 = 0.0000
Log pseudo-likelihood = -4157.416 — Pseudo R2 = 0.0914
(Std. Err. adjusted for 1818 clusters at family level)

Variable	Coefficient	(Std. Err.)	Variable	Coefficient	(Std. Err.)
τ	-0.183	(0.215)	married	0.05	(0.059)
elem	0.201	(0.214)	widowed	-0.192	(0.187)
S2	-0.18	(0.280)	separated	0.011	(0.073)
S3	0.062	(0.281)	divorced	-0.077	(0.087)
S4	-0.044	(0.311)	fulltime	-0.088**	(0.044)
S5	-0.022	(0.282)	nchildren	0.027*	(0.015)
$elem \times \tau$	-0.081	(0.122)	pclabinc	0.032***	(0.008)
$S2 \times \tau$	-0.206	(0.220)	limit	-0.053	(0.046)
$S3 \times \tau$	-0.238	(0.281)	bmi	-0.005*	(0.003)
$S4 \times \tau$	0.025	(0.290)	on leave	-0.113	(0.140)
$S5 \times \tau$	-0.426	(0.280)	unemployed	-0.075	(0.078)
$S2 \times elem$	-0.088	(0.212)	retired	-0.261	(0.168)
$S3 \times elem$	-0.027	(0.322)	disabled	-0.568***	(0.094)
$S4 \times elem$	-0.029	(0.350)	housekeeper	0.035	(0.065)
$S5 \times elem$	-0.514*	(0.308)	student	0.284**	(0.144)
$S3 \times elem \times \tau$	-0.094	(0.345)	stated3	-0.227	(0.253)
$S4 \times elem \times \tau$	-0.206	(0.357)	stated5	0.169	(0.181)
$S5 \times elem \times \tau$	0.583*	(0.317)	stated7	-0.132	(0.281)
$\tau \times female$	-0.067	(0.265)	stated10	-0.12	(0.130)
$elem \times female$	0.135	(0.262)	stated11	-0.182	(0.160)
$S2 \times female$	0.236	(0.223)	stated14	0	(0.193)
$S3 \times female$	0.077	(0.272)	stated15	0.032	(0.211)
$S4 \times female$	0.13	(0.305)	stated16	0.08	(0.209)
$S5 \times female$	-0.141	(0.280)	stated18	-0.246	(0.187)
$elem \times_\tau \times female$	-0.003	(0.157)	stated21	0.037	(0.193)
$S2 \times \tau \times female$	-0.071	(0.268)	stated22	-0.187	(0.212)
$S3 \times \tau \times female$	0.148	(0.355)	stated23	-0.014	(0.189)
$S4 \times \tau \times female$	-0.13	(0.364)	stated25	-0.367**	(0.160)
$S5 \times \tau \times female$	0.367	(0.343)	stated26	0.044	(0.190)
$S2 \times elem \times female$	-0.228	(0.254)	stated31	-0.314*	(0.187)
$S3 \times elem \times female$	-0.317	(0.409)	stated33	-0.136	(0.191)
$S4 \times elem \times female$	-0.548	(0.410)	stated34	0.031	(0.199)
$S5 \times elem \times female$	0.538	(0.395)	stated36	0.011	(0.189)
$S3 \times elem \times \tau \times female$	-0.027	(0.457)	stated38	-0.173	(0.203)
$S4 \times elem \times \tau \times female$	0.371	(0.406)	stated39	-0.104	(0.195)
$S5 \times elem \times \tau \times female$	-0.924**	(0.396)	stated41	-0.305**	(0.152)
jhs	0.067*	(0.039)	stated43	0.113	(0.211)
gender	-0.096	(0.212)	stated44	-0.005	(0.125)
age	-0.002	(0.014)	stated47	-0.1	(0.202)
age^2	0.000	(0.000)	stated48	-0.171	(0.176)
white	0.376***	(0.043)	stated50	0.410*	(0.229)
edu	0.065***	(0.007)	Intercept	0.431	(0.344)

Notes: Significance levels: * = 10%; ** = 5%; *** = 1%.
Variable names: *tau*: time fixed effect; *elem*: group of parent's of elementary school-age children fixed effect (group fixed effect); *Sk*: groups of states k fixed effect (region fixed effect); $elem \times \tau$: group time trend control (group-time interaction); $Sk \times \tau$: group of states' time trend control (region-time interaction); $Sk \times elem$: region-group interaction; $Sk \times elem \times \tau$: triple interaction (region-group-time interaction); all variables of the form $X \times female$ are X variables interacted with the gender dummy $female$; *jhs*: dummy variable equal to one if the individual has at least one children of junior-high-school age; *age*: age in years; age^2: square of age; *white*: white race dummy; *edu*: year of education completed; *married*: married or permanently cohabiting dummy; *widowed*: widowed dummy; *separated*: separated dummy; *divorced*: legally divorced dummy; $fulltime$: equal to one if the individual works less than 36 hours a week during the last year; *nchildren*: number of children (all ages); *pclabinc*: per-capita family labor income in dollars; *limit*: health condition that limits daily activity dummy; *bmi*: body mass index; *onleave*: only temporarily laid off, sick leave or maternity leave dummy; *unemployed*: looking for work, unemployed dummy; *retired*: retired dummy; *disabled*: permanently or temporarily disabled dummy; *housekeeper*: housekeeper dummy; *student*: student dummy; *statedj*: state j fixed effect.

2.7.1 Average Treatment Effects: More details

ATT as a function of missing counterfactuals

We can recover the missing counterfactual $\mathbb{E}(y_{it}^0|elem_i = 1, S_i = 1, \tau_t = 1)$ using equation (2.1), since if we assume that equation (2.3) holds, we have

$$\begin{aligned}\mathbb{E}(y_{it}^0|elem = 1, S = 1, \tau_t = 1) = \mathbb{E}(y_{it}^0|elem = 1, S = 1, \tau_t = 0) \\ + [\mathbb{E}(y_{it}^0|elem = 1, S = 0, \tau_t = 1) - \mathbb{E}(y_{it}^0|elem = 1, S = 0, \tau_t = 0)] \\ + [\mathbb{E}(y_{it}^0|elem = 0, S = 1, \tau_t = 1) - \mathbb{E}(y_{it}^0|elem = 0, S = 1, \tau_t = 0)] \\ - [\mathbb{E}(y_{it}^0|elem = 0, S = 0, \tau_t = 1) - \mathbb{E}(y_{it}^0|elem = 0, S = 0, \tau_t = 0)].\end{aligned} \tag{2.9}$$

We can rewrite the ATT as a function of the unobserved counterfactual $\mathbb{E}(y_{it}^0|elem_i = 1, S_i = 1, \tau_t = 1)$

$$\begin{aligned}ATT = \mathbb{E}(y_{it}^1|elem = 1, S = 1, \tau_t = 1) - \mathbb{E}(y_{it}^0|elem = 1, S = 1, \tau_t = 0) \\ - [\mathbb{E}(y_{it}^0|elem = 1, S = 0, \tau_t = 1) - \mathbb{E}(y_{it}^0|elem = 1, S = 0, \tau_t = 0)] \\ - [\mathbb{E}(y_{it}^0|elem = 0, S = 1, \tau_t = 1) - \mathbb{E}(y_{it}^0|elem = 0, S = 1, \tau_t = 0)] \\ + [\mathbb{E}(y_{it}^0|elem = 0, S = 0, \tau_t = 1) - \mathbb{E}(y_{it}^0|elem = 0, S = 0, \tau_t = 0)].\end{aligned} \tag{2.10}$$

The sample analog of equation (2.10) is the DDD estimator of the ATT

$$\begin{aligned}\widehat{ATT} = (\bar{y}_1^{1,1} - \bar{y}_0^{1,1}) - (\bar{y}_1^{1,0} - \bar{y}_0^{1,0}) \\ - [(\bar{y}_1^{0,1} - \bar{y}_0^{0,1}) - (\bar{y}_1^{0,0} - \bar{y}_0^{0,0})],\end{aligned} \tag{2.11}$$

where $\bar{y}_t^{elem,S}$ is the average of the estimated outcome over individuals in group $elem$, residing in states S, at time t.

ATT in a non-linear model

We rewrite the identifying assumption as follows

$$\begin{aligned}&f^{-1}[\mathbb{E}(u_{it}|elem = 1, S = 1, \tau_t = 1)] - f^{-1}[\mathbb{E}(u_{it}|elem = 1, S = 1, \tau_t = 0)] \\ &- \{[f^{-1}[\mathbb{E}(u_{it}|elem = 1, S = 0, \tau_t = 1)] - f^{-1}[\mathbb{E}(u_{it}|elem = 1, S = 0, \tau_t = 0)]\} \\ &= f^{-1}[\mathbb{E}(u_{it}|elem = 0, S = 1, \tau_t = 1)] - f^{-1}[\mathbb{E}(u_{it}|elem = 0, S = 1, \tau_t = 0)] \\ &- \{f^{-1}[\mathbb{E}(u_{it}|elem = 0, S = 0, \tau_t = 1] - f^{-1}[\mathbb{E}(u_{it}|elem = 0, S = 0, \tau_t = 0)]\}.\end{aligned} \tag{2.12}$$

If equation (2.12) holds, the missing counterfactual is

$$\begin{aligned}\mathbb{E}(y_{it}^0|elem=1,S=1,\tau_t=1) = f\Big\{ & f^{-1}[\mathbb{E}(y_{it}^0|elem=1,S=1,\tau_t=0)] \\ & + \{f^{-1}[\mathbb{E}(y_{it}^0|elem=1,S=0,\tau_t=1)] - f^{-1}[\mathbb{E}(y_{it}^0|elem=1,S=0,\tau_t=0)]\} \\ & + \{f^{-1}[\mathbb{E}(y_{it}^0|elem=0,S=1,\tau_t=1)] - f^{-1}[\mathbb{E}(y_{it}^0|elem=0,S=1,\tau_t=0)]\} \\ & - \{f^{-1}[\mathbb{E}(y_{it}^0|elem=0,S=0,\tau_t=1)] - f^{-1}[\mathbb{E}(y_{it}^0|elem=0,S=0,\tau_t=0)]\}\Big\},\end{aligned} \tag{2.13}$$

and the $ATT = \mathbb{E}(y_{it}^1|elem=1,S=1,\tau_t=1) - \mathbb{E}(y_{it}^0|elem=1,S=1,\tau_t=0)$, can be estimated replacing the expected values by their sample analogs

$$\begin{aligned}\widehat{ATT} = \bar{y}_1^{1,1} - f\Big\{ & f^{-1}(\bar{y}_0^{1,1}) + [f^{-1}(\bar{y}_1^{1,0}) - f^{-1}(\bar{y}_0^{1,0})] \\ & + [f^{-1}(\bar{y}_1^{0,1}) - f^{-1}(\bar{y}_0^{0,1})] - [f^{-1}(\bar{y}_1^{0,0}) - f^{-1}(\bar{y}_0^{0,0})]\Big\},\end{aligned} \tag{2.14}$$

where $\bar{y}_t^{elem,S}$ is the average of the estimated outcome over individuals in group $elem$, residing in states S, at time t.

2.7.2 Robustness: effect of HED on "placebo" treated individuals

To simulate the effect of HED reforms in elementary school on adults without children we need to assign to each individual without children an artificial number of children. Additionally, we need to simulate whether each individual has elementary school-age children and/or of junior-high-school age.

To predict the **number of children** we proceed as follows:

1. Using the sample of parents we estimate by OLS the parameters in the equation

$$nchildren = X\alpha + u,$$

where X includes age, race, gender, marital status, education level, employment status, full-time versus part-time job, total family income level, a set of variables reflecting health status, and state of residence.

2. Using the estimated parameters we predict the number of children in the sample of adults without children

$$\widehat{nchildren} = X\hat{\alpha},$$

3. To obtain an integer number of children we correct the previous estimation

$$\widehat{nchildren} = \begin{cases} 0 & \widehat{nchildren} < 0.5 \\ a & a - 0.5 \leq \widehat{nchildren} < a + 0.5 \text{ for } a = 1, 2, ..., 12. \end{cases}$$

To predict the dummy variable jhs, that is a variable equal to one if the individual has at least one child of **junior-high-school-age**, we proceed as follows:

1. Using the sample of parents we estimate with a probit model the parameters in the

equation

$$Pr(jhs = 1|X) = \Phi(X\gamma),$$

where Φ is the standard normal cumulative density function, and X includes number of children, age, race, gender, marital status, education level, employment status, full-time versus part-time job, total family income level, a set of variables reflecting health status, and state of residence.

2. Using the estimated parameters we predict the probability of having at least one child of junior-high-school age in the sample of adults without children as

$$Pr(\widehat{jhs = 1}|X) = \Phi(X\hat{\gamma}),$$

3. To obtain a binary variable we correct the previous estimation

$$\widehat{jhs}_i = \begin{cases} 1 & Pr(\widehat{jhs_i = 1}|X_i) > \frac{\sum Pr(\widehat{jhs_i=1}|X_i)}{N} \\ 0 & \text{otherwise} \end{cases}$$

Where $\frac{\sum Pr(\widehat{jhs_i=1}|X_i)}{N}$ is a group of states and gender specific average. To predict the dummy variable *elem*, that is a variable equal to one if the individual has at least one child of **elementary-school-age**, we proceed as follows:

1. Using the sample of parents we estimate with a probit model the parameters in the equation

$$Pr(elem = 1|X) = \Phi(X\delta),$$

where Φ is the standard normal cumulative density function, and X includes number of children, indicator of child of high-school-age, age, race, gender, marital status, education level, employment status, full-time versus part-time job, total family income level, a set of variables reflecting health status, and state of residence.

2. Using the estimated parameters we predict the probability of having at least one child of elementary-school-age in the sample of adults without children as

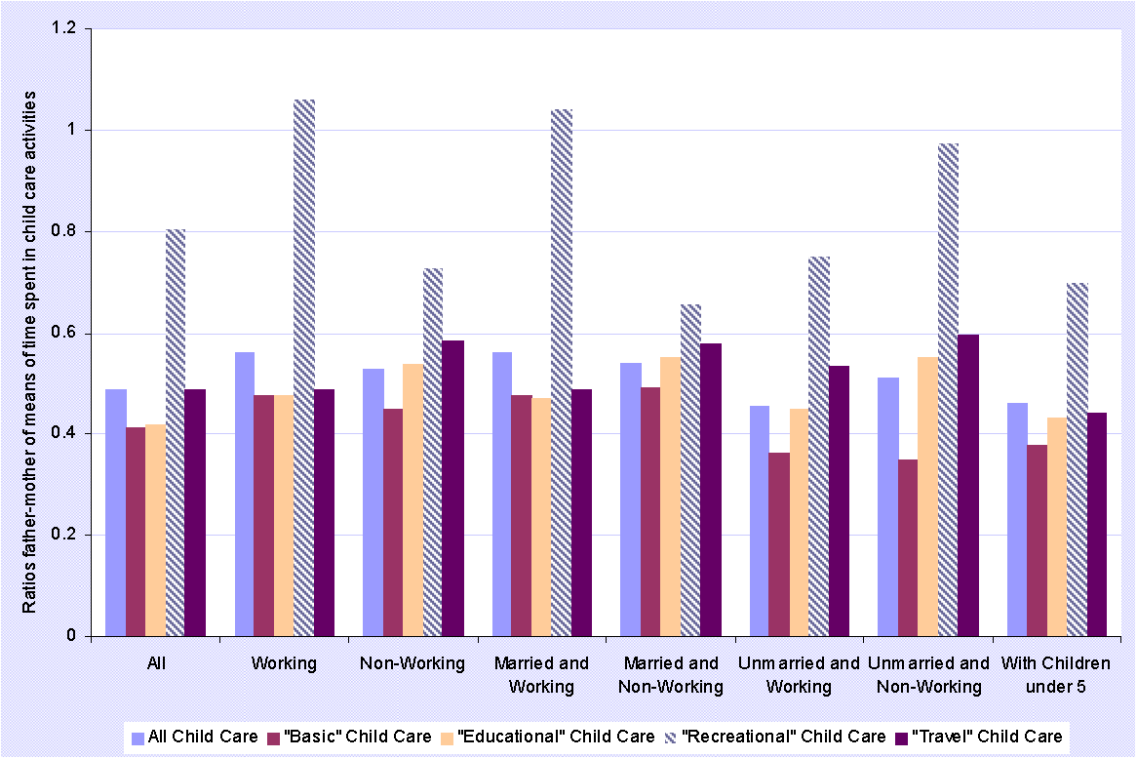
$$Pr(\widehat{elem = 1}|X) = \Phi(X\hat{\delta}),$$

3. To obtain a binary variable we correct the previous estimation

$$\widehat{elem}_i = \begin{cases} 1 & Pr(\widehat{elem_i = 1}|X_i) > \frac{\sum Pr(\widehat{elem_i=1}|X_i)}{N} \\ 0 & \text{otherwise} \end{cases}$$

Where $\frac{\sum Pr(\widehat{elem_i=1}|X_i)}{N}$ is a group of states and gender specific average.

Figure 2.3: Ratios father-mother of means of time spent in childcare activities (hours per week), by demographic subgroups.



Source: Ratios computed using data in Table 1 in Guryan, Hurst, and Kearney (2008) based on the 2003-2006 waves of the American Time Use Survey (ATUS). Childcare activities are classified into: "Basic" childcare (breast feeding, rocking a child to sleep, general feeding, changing diapers, providing medical care to child, grooming child, etc.); "Educational" childcare (reading to children, teaching children, helping children with homework, attending meetings at a child's school, etc.); "Recreational" childcare (playing games with children, playing outdoors with children, attending a child's sporting event or dance recital, going to the zoo with children, taking walks with children, etc.); "Travel" childcare (any travel related to any of the three other categories of childcare). Samples include all individuals between the ages of 21 and 55 (inclusive) who had time diaries summing to a complete day and at least one child under the age of 18.

Chapter 3

Estimating a Dynamic Discrete Choice Model of Health Prevention Decisions: An Application to Flu Vaccination.

Abstract

In this paper I conduct an empirical analysis of the determinants of flu vaccination decisions. Flu vaccination behavior in the adult population (above 65 years old) tends to be highly persistent; additionally, the probability of vaccination increases with age and tends to be higher for individuals with worse health outcomes. To study individual's preventive behavior, I first formulate a stylized life cycle model of prevention decisions using a human capital approach, that highlights the importance of the dynamic dimension in these decisions. The main aspects of the model are: i) Influenza immunization is a health investment, which affects the evolution of future health stock and hence, affects individuals' future utility. This investment implies some monetary and non-monetary costs; ii) Vaccination has higher returns for individuals with health conditions that increase the risk of influenza-related complications; iii) Experience with the vaccine in the previous period reduces the cost of current prevention effort, generating habit persistence. I estimate a reduced-form of the demand function of vaccination implied by the model using dynamic probit models, that allow me to disentangle how much of the observed persistence in vaccination decisions are due to state dependence (habit persistence), unobserved heterogeneity, and health risks or other observable characteristics. I also analyze whether individuals' incentives to pursue prevention change through the life cycle. I use data from the Medicare Current Beneficiary Survey for the period 2001-2004. Results suggest that the three sources −state dependence, unobserved heterogeneity, and health risks and other individual characteristics− play a role in explaining the persistence in vaccination decisions. However, health risks and individual characteristics have a lower effect once state dependence and unobserved heterogeneity are

taken into account. The results also show that the incentives to vaccinate change with age but not with self-assessed health status.

3.1 Introduction

Influenza is an infectious disease that can have severe health consequences for the adult population, pneumonia being the most frequent complication. Influenza and pneumonia are the sixth cause of death for the population above 65 years old in the US, with a mortality rate of 140 per 100,000 inhabitants.[1] According to the Agency for Health Care Research and Quality, during the period 1997-2006 these illnesses accounted for 6% of total hospital stays for the elderly. Each hospital stay implied an average cost of 9,500 dollars for influenza and above 14,000 dollars for pneumonia.[2] Hence, influenza is an important public health concern.

Given that influenza immunizations have been demonstrated to be cost-effective for persons aged 65 and older (Maciosek, Solberg, Coffield, Edwards, and Goodman, 2006), various public health organizations, including the World Health Organization, recommend annual vaccination for the elderly (Stohr, 2003). Persons with chronic health conditions who face a higher risk of influenza-related complications, such as heart and lung diseases, are particularly encouraged to vaccinate. In the US, even though Medicare subsidizes the annual influenza immunization for its beneficiaries, the coverage rate only reached 65% of the population 65 years old and older in 2000 and this percentage remained almost constant since then.[3] Increasing vaccination coverage is one of the objectives in the agenda of the U.S. Department of Health and Human Services. Reaching a vaccination coverage of 90% by 2010 was an unmet objective of the Healthy People initiative but it is still the target to be achieved by 2020.[4]

In this paper I conduct an empirical analysis of the determinants of individuals' vaccination decisions. First, I formulate a stylized life cycle model of primary prevention decisions using a human capital approach.[5] I estimate a reduced-form model of the demand function of

[1] According to data from the National Center for Health Statistics, Trends in Health and Aging, for the year 2004.

[2] The average length of stay due to influenza in the period 1997-2006 was 4.7 days for people aged 65-84 and 5.8 for those aged 85 and older. For pneumonia the average length of stay was 6.5 for people 65 years old and older. See more details on hospitalizations in Table 3.11 in Appendix 3.7.1.

[3] Medicare part B covers both the costs of the vaccine and its administration by recognized providers. Medicare part A does not cover this benefit, but only a small share of the Medicaid population is covered by part A alone. The data for vaccination coverage comes from the National Center for Health Statistics (National Health Interview Survey, sample adult questionnaire).

[4] *"Healthy People 2010 was an initiative carried out by the U.S. Department of Health and Human Services, who set a comprehensive nationwide health promotion and disease prevention agenda designed to enhance population health through preventive behaviors. Healthy People 2010 contained 467 objectives designed to serve as a framework for improving the health of all people in the United States during the first decade of the 21st century"*(Healthy People 2010 Database, NCHS). Healthy People 2020 sets the objectives to be achived by 2020.

[5] Preventive measures can be classified according to their effects on health (Kenkel, 2000), and vaccination

vaccination implied by the model. I use dynamic panel probit models and I contrast the results with the predictions of the theoretical dynamic model. Using panel data from the Medicare Current Beneficiary Survey (MCBS) for the period 2001-2004, I disentangle the effects of state dependence (habit persistence), unobserved heterogeneity, and health risks and other time-invariant observable characteristics, on the probability of vaccination.[6] I also analyze the effect of age and self-assessed health status on vaccination decisions, as well as the effect of other individual characteristics that change over time.

In order to capture the main features of vaccination decisions I construct a simple dynamic discrete choice model of prevention decisions. In the model, individuals face the following trade-off. Prevention generates a benefit in terms of better health in the long run at a cost today, for instance, exerting some effort −time, search for information− or incurring some monetary costs. Individuals with worse health status may suffer greater health losses when getting the flu, which implies that prevention has higher returns for them. Additionally, the model allows for current immunizations decisions to have a direct effect on the cost of future decisions, by assuming that past experience with the vaccine reduces the cost of current prevention effort. For example, for the case of Medicare, Parente, Salkever, and DaVanzo (2005) argue that through the experience with the vaccine individuals learn that vaccination is a benefit fully covered for Medicare beneficiaries, i.e., it is free of charge.

In the empirical analysis I use the approaches proposed by Heckman (1981) and Wooldridge (2005) to deal with the endogeneity problem generated in this type of dynamic models by the presence of the lagged dependent variable as a regressor, together with unobserved heterogeneity (usually called in the literature the "initial condition problem"). Additionally, I compare the results of the dynamic models estimated with the Wooldridge and Heckamn approaches with models that ignore state dependence, which I refer to as static models.

Understanding the determinants of individuals' preventive behavior is important to improve the design of policies aimed to increase flu vaccination coverage. Previous empirical works have studied how health status (Mullahy, 1999; Wu, 2003), consumer knowledge (Parente, Salkever, and DaVanzo, 2005), physician quality (Schmitz and Wubker, 2010; Maurer, 2009), and perceived risks (Mullahy, 1999; Ayyagari, 2007), affect vaccination decisions of the older population. My paper builds upon this literature and analyze the determinants of vaccination in a dynamic setting. Most of the analysis in the previous literature have been carried out using cross-sectional data. A complete theoretical framework incorporating prior experience with the vaccine was considered in the work by Mullahy (1999), but due to data limitations prior experience was not introduced in the analysis. Although

belongs to what is called *primary prevention*. Primary preventive measures allow to reduce the probability of occurrence of a disease. This category comprises also public sanitation policies and individual lifestyles (as regular exercise and non-smoking).

[6]The dependence of current decision on lagged ones is known in the empirical literature as "true state dependence" (Heckman, 1981).

Parente, Salkever, and DaVanzo (2005) consider state dependence in their empirical analysis, they do not deal with the potential endogeneity problems introduced by the lagged dependent variable.[7] Compared to the previous literature, the use of panel data allows to test not only the importance of health conditions and other individual characteristics on the vaccine demand, but also to disentangle the relative importance of state dependence and unobserved heterogeneity. As I show in Section 3.3, individual vaccination decisions are highly persistent over time. The nature of this persistence is important for policy design purposes. The presence of state dependence, for instance, implies that the any public campaign that induces individuals to get the vaccine for the first time, will have persistent effects in subsequent periods, while this would not be the case if persistence is due to unobserved time-invariant characteristics.

The results suggest that the three possible sources of persistence (state dependence, unobserved heterogeneity, and health risks and other time-invariant individual characteristics) do play a role in explaining it. The Heckman model provides a better fit to the observed data than the Wooldridge approach, hence it is the preferred specification. I find that individuals who get the vaccine in a given year are, on average, between 12 to 14 percentage points more likely to get the vaccine in the next year than those who did not get it. Also, individuals' choices depend to a large extent on unobserved heterogeneity, which accounts for 60% to 80% −depending on the estimation strategy− of the total variance of the error term. Preexisting chronic conditions, such as diabetes or arthritis rheumathoid, increase the probabilty of vacicnation by 4 percentage points. Other socioeconomic characteristics that are constant over time, such as more education or being white increase the probabilty of vacicnation by 5 percentage points.

I also find that individual's behavior adjusts to changes in the perceived risks of influenza-related complications, which are not necessary constant over time. For example, I find that individuals do increase their likelihood of vaccination if in the previous period they experienced a health shock (a new episode of respiratory illness or stroke). Also, as in previous literature, I find that married individuals, white, and with supplemental private or public health insurance are more likely to engage in vaccination. I show that the incentives to vaccinate change with age although it does not change with self-assessed health status, as opposed to the findings in Wu (2003) and Mullahy (1999). Vaccination propensities tend to increase with age but showing a slight slowdown at advanced ages. Finally, the comparison of the preferred dynamic and static specifications show that conclusions do not change qualitatively when controlling for state dependence, but the dynamic model fit the data better to predict the sequence of vaccination decisions.

[7] Parente, Salkever, and DaVanzo (2005) point out that prior vaccination use increases consumer knowledge about Medicare benefits through experience and hence, it makes individuals more likely to get the vaccine in the next year. Also, individuals without previous experience may also underestimate the risk of exposure to the illness, the severity of the disease, or the effectiveness of the vaccine to prevent the disease.

The implications of these results for public health policy purposes are, at least, twofold. First, given that individuals do internalize the fact that certain health risks increases the benefit of vaccination, this channel can be exploited to increase even further vaccination take-up rates. For instance, public campaigns alerting the population that influenza-related complications are more acute for individuals with certain health conditions, would increase vaccination coverage. Second, any public campaign that induces individuals to get vaccinated for the first time, will have effects on subsequent periods through the habit persistence channel.

The paper is organized as follows. Section 3.2 presents the conceptual framework that guides the empirical application. Section 3.3 describes the data and the sample selection. Section 3.4 describes the empirical strategy I follow for the reduced form estimation, as well as the econometric issues related to the estimation of dynamic probit models. Section 3.5 presents the results and Section 3.6 concludes.

3.2 Theoretical Framework

In this section I present a model of vaccination decisions which is characterized by two main features. First, the time dimension is introduced in the problem solved by the individual, i.e., decisions are made in a life-cycle context. The second feature of the model is the introduction of uncertainty, since the evolution of the health stock is stochastic. Vaccination constitutes a means to reduce uncertainty, as it reduces the probability of occurrence of a particular illness.

In the model, individuals maximize the present value of their lifetime utility. Own health stock is a consumption good, i.e., it enters in the utility function, as well as human capital that can be modified through individual actions. I assume that health stock has a stochastic component that accounts for the uncertain evolution of health. Every period, individuals are exposed to the occurrence of a negative health shock, the flu. If the individual receives the negative health shock, there is positive probability of recovering from it and maintaining his health stock unchanged. Nevertheless, the shock may produce a deterioration in the health stock that remains for more than one period.[8] Individuals are concerned with the magnitude of these effects because it is in their interest to increase the probability of being in good health in the future.

Individuals may affect the transition probabilities between health states using prevention methods that reduce the probability of occurrence of the negative shocks, acting as "self-insurance". Under this setting, the dual role of health as a consumption good and a human capital is key in determining individual's incentives to pursue prevention.

[8]Long lasting effects of influenza and pneumonia are more likely to occur among the elder individuals and among those with chronic conditions because of the high risk of complications due to the flu.

In the model, individuals face a trade-off between the long-run benefits of using prevention measures, i.e., the higher expected health stock in the future which increases the present value of expected utility, and the current monetary and non-monetary costs associated with these measures. In this model, the experience with the vaccine in the previous period reduces current non-monetary costs of prevention.

Timing. Individuals have a finite life time $t = 1, ..., T$. However, it is possible that individuals die before time period T. At the beginning of each period an individual decides whether to pursue primary prevention, given his health stock, h_t, his previous prevention decision, d_{t-1}, and other individual characteristics, w_t, in order to reduce the probability of receiving a negative health shock that may occur at the end of the period.[9] Prevention reduces the probability of occurrence of a negative health shock and, as a consequence, increases the probability of enjoying better health in the next period. Prevention also increases the probability of surviving. The decision variable is denoted by d_t and it may take two values $d_t \in \{0, 1\}$ indicating whether the individual gets the vaccine. Every period, only one health shock may occur, $s_t \in \{0, 1\}$, and its realization is unknown at the moment the prevention decision is made. If death occurs it is assumed to happen at the end of the period.

Preferences. The current period utility function is modeled as an additive separable random utility given by the following expression:

$$U_t(h_t, d_t, d_{t-1}, w_t, \epsilon_t) = \begin{cases} u(h_t, w_t) + \epsilon_{0,t} & \text{if} \quad d_t = 0 \\ u(h_t, w_t) + \epsilon_{1,t} - C_p(d_{t-1}, h_t, w_t) & if \quad d_t = 1 \end{cases}$$

where $C_p(1, h_t, w_t) < C_p(0, h_t, w_t)$.

The current health stock, h_t, enters the utility function because health is perceived as a consumption good; current prevention effort, $d_t = 1$, generates desutility in the current period, C_p, which reflects monetary and non-monetary costs of prevention. These costs depend on individual health, h_t, other individual characteristics, w_t, and past experience, d_{t-1}. I consider that part of non-monetary costs are related to knowledge referred to the vaccine uptake and its characteristics, and they are reduced if the individual experienced with the vaccine in the past.[10] Finally, utility is affected by an idiosyncratic choice-specific preference shifter, ϵ_{jt}.

[9] I omit in this section the use of the subscript i to refer to individuals.

[10] Notice that part of these informational costs may be a fixed cost that would disappear the **first** time the vaccine is consumed. I do not model this possibility because in the data set there is no information about the time in which individuals got the vaccine for the first time.

Health production. The health of each individual is assumed to evolve stochastically as a function of current health, h_t, health investments, d_t, and individual characteristics, w_t. Prevention activities are the only form of health investments. I assume that there is a finite number, H, of health states. For simplicity, I assume that there are only two possible health states, good, h_g, and bad, h_b.

Uncertainty about future health outcomes is modeled in the following way: first, individuals face the possibility of contracting the illness, $s_t = 1$. The probability of contracting the illness, $S(h_t, w_t, d_t)$, is a function of the individual's current health stock, h_t, individual characteristics, w_t, and current prevention decision, d_t. By assumption this probability is always greater than zero and it decreases if the individual gets the vaccine, that is $S(h_t, w_t, 0) > S(h_t, w_t, 1) > 0$. Also, it depends negatively on the level of health stock, that is $S(h_b, w_t, d_t) > S(h_g, w_t, d_t)$.

The health stock evolves according to a first order Markov process with transition matrix Π. These transitions are conditional upon survival and depend on current health stock, h_t, the realization of the shock, s_t, and individual characteristics, w_t. Once the shock is realized, health transition probabilities do not depend on d_t. Each element of the transition matrix is denoted by $\pi_{ml}(w_t, s_t)$, which correspond to the probability of being in the next period in the health state h_m given that current health state is h_l, for m and $l \in \{g, b\}$, individual characteristics are w_t and the realization of the shock is s_t. If the individual contracts the illness ($s_t = 1$) I assume that, other things equal, an individual with bad health, h_b, is less likely to be in good health than someone with good health, h_g. This is to say, $\pi_{gl}(w_t, s_t) < \pi_{gg}(w_t, s_t)$, for $l \in \{g, b\}$.

At the moment the individual takes the prevention decision, the negative health shock has not been realized, which implies that when solving his dynamic problem he needs to integrate out the health shock. Then, health transition probability at the moment the decision is made can be written as:

$$F(h_m|h_l, w_t, d_t) = S(h_l, w_t, d_t)\pi_{ml}(w_t, 1) + (1 - S(h_l, w_t, d_t))\pi_{ml}(w_t, 0), \quad \text{for } m \text{ and } l \in \{g, b\}. \tag{3.1}$$

Given previous assumptions, we have that $F(h_g|h_l, w_t, 1) > F(h_g|h_l, w_t, 0)$ for $l \in \{g, b\}$.

Survival probability. The probability that an individual survives until the end of the period is denoted by $p_{su}(h_t, w_t, d_t)$. I assume that the probability of survival increases with the use of the vaccine but it decreases with age, which is one of the variables in the vector w_t. Also, I assume that the ratio $\frac{p_{su}(h_t,w_t,1)}{p_{su}(h_t,w_t,0)}$, which measures the effectiveness of the vaccine in terms of extending life, decreases as age increases, $t \to T$, but this ratio is always higher for individuals in bad health status.

Maximization problem. Individuals maximize their lifetime discounted utility by making

sequential choices over health prevention, d_t, in each time period. Individuals are forward looking. In any period τ the individual solves the following maximization problem:

$$\max_{d_t \in \{0,1\}_{t=\tau}^{t=T}} \mathbb{E}_\tau [U_\tau(h_\tau, d_\tau, d_{\tau-1}, w_\tau, \epsilon_\tau) + \sum_{t=\tau+1}^{T} \beta^{t-\tau} p_{su}(h_{t-1}, w_{t-1}, d_{t-1}) U_t(h_t, d_t, d_{t-1}, w_t, \epsilon_t) | h_\tau, w_\tau, d_{\tau-1}] \quad (3.2)$$

We can rewrite this problem as a dynamic programming problem. At any period $t = 1, ..., T$ the problem for any individual consists on maximizing the expected present value of the remaining lifetime rewards. Lets define $\mathbf{\Omega}_t = (h_t, w_t, d_{t-1}, \epsilon_t)$ as the vector of state variables. Then, the maximum expected present value of lifetime utility at time t given $\mathbf{\Omega}_t$ is:

$$V_t(\mathbf{\Omega}_t) = \max_{\mathrm{d}_j \in \{0,1\}} \{ U(h_t, d_j, d_{t-1}, w_t, \epsilon_t) + p_s(h_t, w_t, d_j) \, \beta \mathbb{E}(V_{t+1}(\mathbf{\Omega}_{t+1}) | \, h_t, d_j, w_t) \} \quad (3.3)$$

I define the alternative-specific value function at time t as $V_t^j(\mathbf{\Omega}_t)$, for the alternatives of not doing prevention ($j = 0$) and doing prevention ($j = 1$). Every period, an individual compares the present discounted expected utility over the remaining lifetime from doing prevention today and making optimal decisions in the future, $V_t^1(\mathbf{\Omega}_t)$, with the respective discounted value of not doing prevention and making optimal decisions in the future, $V_t^0(\mathbf{\Omega}_t)$, that is $V_t(\mathbf{\Omega}_t) = \max \left\{ V_t^0(\mathbf{\Omega}_t), V_t^1(\mathbf{\Omega}_t) \right\}$. He decides to pursue prevention if and only if $V_t^1(\mathbf{\Omega}_t) \geq V_t^0(\mathbf{\Omega}_t)$. That is,

$$\begin{aligned} V_t^1(\mathbf{\Omega}_t) &\geq V_t^0(\mathbf{\Omega}_t) \\ \mathbf{v}_t^* - C_p(d_{t-1}, h_t, w_t) + u_t &\geq 0 \end{aligned} \quad (3.4)$$

where

$$\mathbf{v}_t^* = p_s(h_t, w_t, 1) \, \beta \mathbb{E} \left[V_{t+1}(\mathbf{\Omega}_{t+1}) | \, h_t, 1, w_t \right] - p_s(h_t, w_t, 0) \, \beta \mathbb{E} \left[V_{t+1}(\mathbf{\Omega}_{t+1}) | \, h_t, 0, w_t \right]$$

and $u_t = \epsilon_t^1 - \epsilon_t^0$.

In the Appendix 3.7.2 I solve the model under a particular set of assumptions and I present a numerical example. The implications derived from it are the following. First, if previous experience with the vaccine reduces the current cost of prevention, then individuals that get the vaccine in one period are more likely to do it again in the following period, compare with those who do not get the vaccine. Second, the probability to get the vaccine increases with age, for a given health status, although there is a slowdown at the end of life. Third, only at advanced ages individuals with worse health status are more likely to get the vaccine

relative to those in better health. The gap increases with age and then closes at the end of life. At younger ages, it could be the case that individuals in better health are more likely to get the vaccine than individuals in worse health. This could happen because although the expected gains from vaccination in terms of health are lower than for individuals in worse health, life expectancy is higher so they have a longer time horizon to enjoy the health gains from vaccination.

3.3 Data and sample selection

3.3.1 Data description

I use annual data from the Access to Care Files of the Medicare Current Beneficiary Survey (MCBS) for the years 2001 to 2004. Respondents for the MCBS are sampled to be representative of Medicare population as a whole. The MCBS is a longitudinal survey where sampled individuals are interviewed during four years.

This dataset collects survey's information on a broad spectrum of individual's health and socioeconomic characteristics, as well as health related behavior. Additionally, it collects information about access to care, insurance coverage, financial resources, and potential family support.

The Access to Care files sample the "always enrolled" Medicare population, which consists of those enrolled in one part of Medicare, Part A or B, or enrolled in both parts on January 1 of that year and who remain enrolled through the end of December.

The MCBS survey's data is matched with administrative records, that register the individuals' claims with a detailed description of the use of health care services during a year. The claims record the utilization of services rendered and reimbursed under fee-for services during a calendar year. The services that Medicare beneficiaries have access to are of 7 different types: inpatient hospital stays, skilled nursing facility, home health care, hospice care, outpatient services, physician's services, and durable medical equipment. Any individual may use any of these services during a calendar year, and each time a service is used it is registered as a claim record. Each claim registers the diagnosis the patient presents each time he uses health care services, according to the International Classification of Diseases, 9th Revision, Clinical Modification medical codes (ICD-9), and the date in which the service is used.

I consider a balanced sample of respondents who were not living in facilities, such as nursing or retirement homes, during the whole period under analysis.[11] The sample is also restricted

[11]Given that I am dealing with an old population, some attrition is expected due to death, serious illnesses, or people moving to institutionalized care. So, people who remain in the balanced panel are more likely to be healthier. This may affect the estimated effect of the measures of health status on vaccination decisions, but the direction of the bias is not clear. On the one hand, as healthier individuals remain in the balanced sample, the effect of the health measures I control for (heart problems, stroke, chronic lung disease, respiratory illnesses, self reported health status, change in health relative to previous year) on the

to individuals aged 65 and over giving a pooled sample size of 5313 observations.

The outcome of interest is whether the respondent had the flu vaccine during the influenza season. Individuals are asked whether they had a influenza shot during the last winter. Given the data available, I am able to construct a binary indicator of vaccination for the 2000-2001, 2001-2002, 2002-2003 and 2003-2004 flu seasons. Flu seasons run from October to April in the US and outbreaks are more likely to occur from late December to early March (ACIP, 2008). The optimal time for immunization of high-risk groups is recommended between October and November.

To account for the multiple dimensions of health, I consider different measures of health status available from the MCBS. I give a possible interpretation of each of these variables in the context of the theoretical model and I define the expected effects of these variables on the vaccination decision.[12]

Risk factors. As an overall measure of health I use self-reported health status. The survey ask individuals to assess their health, explicitly indicating to compare it with respect to individuals of the same age. I construct three categories (Good, Regular, and Bad) and interact these variables with age. According to the predictions of the theoretical model, individuals' incentives to pursue prevention change through the life cycle and they may differ for individuals with different perceived health status. In particular, vaccination propensity should increase with age, because age is a risk factor, but it may slowdown at advanced ages, because the planning horizon is shorter. Additionally, the gap of vaccination propensities between individuals with worse health and better health may increase as individuals get older.

A second set of variables are indicators of diseases that are recognized to increase the risk of influenza-related complications, such as cancer, heart disease, diabetes, chronic lung disease, stroke, and arthritis rheumatoid. I construct indicators that take the value one if the individual reports to had suffered some of these diseases prior to the beginning of the sample period. Hence, these are preexisting health conditions for the window of time analyzed and all of them are expected to increase the likelihood of taking the vaccine.

Finally, a third set of risk factors are indicators of recent health events (health "shocks") that increase the risk of influenza related complications. I include two variables that indicate if new events of heart disease or new events of stroke occurred in the previous year. These indicators are based on answers given to the questions in the survey.

Although I do not modeled it in the theoretical framework, I additionally include an indicator of the occurrence of respiratory illnesses during the period prior to the vaccination

probability of taking the vaccine could be *underestimated*. Those who remain in the sample are healthier and, hence, less likely to take the vaccine than those that are dropped from the sample. On the other hand, individuals not in the sample could be those less likely to pursue health preventive activities and, hence, more likely to suffer health shocks. If this is the case, the effect of health outcomes could be *overestimated*.

[12] See the appendix 3.7.4 for a detailed description of the variables.

decision. As pointed out by Ayyagari (2007) and Mullahy (1999) individuals may associate this shock to an increase in the risk of getting the flu, which in turn increases their incentives to get the vaccine in the next period. To construct the indicator that refers to the occurrence of respiratory illness I use administrative information on Medicare claims. I construct a binary variable which takes the value 1 if during a given period any disease, coded as influenza, pneumonia, or other respiratory disease, is present in the diagnosis of any of the individual's claims, and 0 otherwise. I am able to construct health shock variables for the periods prior to the flu seasons 2001-2002, 2002-2003 and 2003-2004.[13]

Physical limitations. Individuals with physical limitations may find more costly to get vaccinated. For instance, they may require the company of other person to go to the place to get the flu shot. I consider three variables that may capture physical limitations: a binary variable that indicates whether the individual had a broken hip last year and a variable that counts the number of limitations with activities of daily living (ADLs)

Planning horizon. There is empirical evidence (Benitez-Silva and Ni, 2008) of the positive relation between self-reported health changes and expected longevity. Hence, I use an indicator that takes the value one if the individual reports that his or her health is worse or much worse than in the previous year as a proxy for expected longevity. According to the theoretical model, individuals reporting that their health is worse relative to previous year may be less likely, ceteris paribus, to get the vaccine, as the time horizon to enjoy the benefits of the vaccine is shorter.

Individual characteristics. Socioeconomic and demographic characteristics may be also important factors determining preventive behavior. Through all the analysis I consider marital status, gender, race, and education. Education is a particularly important factor since there is evidence that more educated individuals have a greater knowledge of health related issues, and therefore, they may be more likely to pursue health preventive activities (Kenkel, 1991; Cutler and Lleras-Muney, 2006; Park and Kang, 2008; Parente, Salkever, and DaVanzo, 2005). I also consider whether individuals have an additional source of health insurance coverage (public or private) that complements Medicare benefits or whether individuals have Medicare part A only and, as a consequence, free vaccination is not available for them.

[13]Influenza, pneumonia and other respiratory illness shocks are identified using the diagnosis codes based ICD-9 medical codes which are 487, 480-486 and 460-519, respectively. The construction of health shock variables from Medicare claims data relies on the assumption that any disease the individual had (sufficiently acute to demand medical services) should be captured by Medicare claims. Specifically, I am assuming that individuals do not resort to medical services outside the Medicare orbit in the event of illness and that diagnosis done by doctors are correct. An individual who did not use any Medicare service during the period is considered as not suffering any disease during it.

3.3.2 Summary statistics

A salient characteristic of individual vaccination decisions that motivates the study of vaccination decisions in a dynamic framework is that they are highly persistent over time. Table 3.1 shows that flu vaccination has a clear persistent pattern. Approximately 95% of individuals who get the vaccine in a given year are expected to get it again in the following year. Also, 76% of individuals who do not vaccinate in a given year are likely to continue with the same behavior in the next period.

Medicare part B covers both the costs of the vaccine and its administration by recognized providers. Hence, all Medicare part B beneficiaries can potentially get the vaccine for free every year. One way through which Medicare beneficiaries learn about this benefit and the need of receiving a flu shot, is through an annual guide send by the Centers for Medicare & Medicaid Services (CMS) that contains information about health plans, prescription drug plans, and rights and protections. In particular, beneficiaries receive information indicating which preventive services are covered. This could in part explain the persistent behavior. However, this information does not reach all beneficiaries since the most common reason beneficiaries gave for failing to get a flu shot was that they did not know that the federal government recommends and pays for it (B. Baker, 1999). Other reason could be that individuals with health problems receive systematically advice of their doctors to receive the shot. A third reason, as suggested by Parente, Salkever, and DaVanzo (2005), is that as beneficiaries experience once with the vaccine they learn that vaccination is a benefit fully covered for Medicare beneficiaries.

Table 3.1: Transition rates. Individuals aged 65+, living in the community.

Year t status	Year $t+1$ status	2001-2002	2002-2003	2003-2004
No vaccination	No vaccination	0.765	0.796	0.761
	Vaccination	0.235	0.204	0.239
Vaccination	No vaccination	0.054	0.055	0.047
	Vaccination	0.946	0.945	0.953
N		2,325	2,325	2,325

Source: MCBS.

Table 3.2 presents summary statistics on a selection of relevant socioeconomic and health status variables for the full sample in the year 2001 and for subgroups classified according to their vaccination sequence during the period 2001-2004. There are 16 possible decision paths, but I restrict to five mutually exclusive cases: the subsample of individuals who get the vaccine each year of the sample period (column 3); individuals who skip the vaccine each year (column 2); individuals who experiment a single transition from vaccination to no vaccination (column 4), that is, sequences '1000', '1100' and '1110'; individuals that experiment a single transition from no vaccination to vaccination (column 5), with sequences

'0001', '0011' and '0111'; and individuals that experiment multiple transitions (column 6). To asses if the different sequences of vaccination are explained by observed characteristics, columns (2) to (6) of Table 3.2 show summary statistics for the five subsamples classified according to the sequence of decisions. The differences among subsamples can be summarized as follows. Regarding the socioeconomic characteristics and comparing to the average, individuals that engage in prevention every year −column (3)− are older, better educated, and are more likely to be white, married, and to have supplemental health insurance. On the contrary, individuals who never engage in prevention −column (2)− are more likely to be younger and less educated, and they are less likely to have a supplemental health insurance, than the average.[14]

Health differences are also well established between these two extreme groups: individuals pursuing prevention each year of the sample period have in general worse health than the average, while individuals who skip prevention are in better health than average. Indeed, almost all measures of health status at the beginning of the sample period are worse for the always takers than for the never takers.[15] This could be in part explained by a supply effect, i.e., doctors recommend their patients to take the flu shot.

Individuals that present a single transition from vaccination to no vaccination −column (4)− are less educated and more likely to be white and married. Individuals in this group are more likely to have antecedents of heart attack and more likely to have had a broken hip last year and to have more limitations with activities of daily living.[16] Individuals experiencing a single transition from no vaccination to vaccination −column (5)− are younger, more likely to have supplemental insurance, but there is no a clear pattern about their health.

The patterns that arise when comparing columns (2) and (3) of Table 3.2 are consistent with the determinants of prevention suggested by some of the the existing empirical literature: 1) Health seems to affect individual incentives to get the vaccine; 2) Socioeconomic variables −education, race, marital status, and supplemental insurance coverage− seem to play an important role on individual behavior as well.

3.4 Empirical strategy

I estimate a reduced-form demand function of preventive care, d_t, implied by equation (3.4), which indicates that the individual will pursue prevention if $\mathbf{v}_t^* - C_p(d_{t-1}, h_t, w_t) + u_t \geq 0$. Let subscript i denote individual observations and let's assume that the cost associated to prevention is $C_{p,it} = \varsigma_{it} - \gamma d_{i,t-1}$. The cost of prevention is assumed to depend on individual characteristics, captured in the term ς_{it}. I make the simplifying assumption that

[14] In all cases the differences are statistically significant.

[15] The differences between theses two groups are statistically significant for self-reported health status, hypertension, angina, other health problems, and diabetes.

[16] In all cases the differences are statistically significant.

Table 3.2: Descriptive Statistics, year 2001

		Sequence of vaccination decisions 2001-2004				
	Full Sample (1)	Never have Vaccine (2)	Always have Vaccine (3)	Single transition from vacc. to non-vacc. (4)	Single transition from no vacc. to Vacc. (5)	Multiple Transitions (6)
Individual Characteristics						
Age	74.63	73.49	75.38	75.55	73.23	73.56
Female	0.57	0.53	0.56	0.64	0.57	0.63
White	0.87	0.85	0.91	0.62	0.80	0.74
Education	0.69	0.63	0.72	0.57	0.66	0.62
Married	0.55	0.51	0.58	0.42	0.54	0.49
Insurance coverage						
Medicare Part A Only	0.03	0.04	0.02	0.06	0.04	0.04
Supplemental Insurance	0.92	0.87	0.93	0.89	0.96	0.87
Health measures						
Health Good	0.48	0.55	0.48	0.43	0.43	0.43
Health Regular	0.32	0.29	0.32	0.28	0.34	0.34
Health Bad	0.20	0.16	0.19	0.28	0.23	0.23
Hypertension	0.53	0.44	0.56	0.49	0.53	0.54
Heart Attack	0.12	0.10	0.13	0.21	0.10	0.15
Angina	0.09	0.07	0.11	0.08	0.06	0.05
Other Heart Problems	0.27	0.19	0.30	0.23	0.27	0.24
Stroke	0.09	0.08	0.09	0.08	0.12	0.11
Cancer	0.16	0.13	0.17	0.09	0.15	0.19
Chronic Lung Disease	0.11	0.07	0.12	0.17	0.09	0.12
Diabetes	0.17	0.12	0.17	0.17	0.19	0.20
Arthritis Rheu	0.08	0.07	0.08	0.06	0.12	0.08
Broken Hip	0.02	0.04	0.01	0.04	0.04	0.03
Number of ADLS	0.47	0.43	0.44	0.83	0.59	0.50
N	1771	339 *(19%)*	1039 *(59%)*	53 *(3%)*	223 *(13%)*	117 *(7%)*

the effect of previous experience, $d_{i,t-1}$, affects all individuals in the same way and generates a reduction of prevention costs. The dynamic discrete choice equation for individual i in period t is then:

$$d_{it} = \begin{cases} 1 & \text{if} \quad \mathbf{v}_{it}^* - \varsigma_{it} + \gamma d_{i,t-1} + u_{it} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

The variable $\mathbf{v}_{it}^*$ can be considered a latent variable representing the expected increment to future utility for individual i if he vaccinates in period t. To parameterize the reduced-form model, I assume $\mathbf{v}_{it}^* - \varsigma_{it}$ follows a linear model in parameters:[17]

$$\mathbf{v}_{it}^* - \varsigma_{it} = x_{it}\beta + c_i, \tag{3.6}$$

where x_{it} is a vector of covariates, including individual characteristics, measures of health status, year, and geographical dummies, and c_i captures individual unobserved characteristics that remain unchanged through time. Replacing (3.6) into (3.5) we obtain the following latent variable version of the demand for prevention:

$$d_{it}^* \quad = \quad x_{it}\beta + \gamma d_{i,t-1} + v_{it} \tag{3.7}$$

The lagged decision, $d_{i,t-1}$, is a binary variable which takes value 1 if the individual was immunized in the previous year. In the dataset I observe the actual decisions, d_{it}, therefore, the estimated model is the following:

$$d_{it} \quad = \quad \mathbf{1}(d_{it}^* > 0) = \mathbf{1}(x_{it}\beta + \gamma d_{i,t-1} + v_{it} > 0) \tag{3.8}$$

The error term v_{it} has the following structure:

$$v_{it} \quad = \quad c_i + u_{it} \tag{3.9}$$

$$u_{it} \quad \sim \quad iid \;\; Normal(0,1) \tag{3.10}$$

Although the errors u_{it} are assumed serially independent, the composite error term, $v_{it} = c_i + u_{it}$, will be correlated over time due to the individual-specific time invariant term, c_i. The specific form of unobserved heterogeneity assumed, i.e., additive, individual-specific, and time invariant, implies equi-correlation between the v_{it} component in any two different periods:

[17] Under this setting ς is the parameter associated with the current monetary and non-monetary costs of prevention.

$$\rho = Corr(v_{it}, v_{is}) = \frac{\sigma_c^2}{1+\sigma_c^2} \qquad t, s = 2, ..., T; t \neq s \tag{3.11}$$

Note that the cross-period correlation, ρ, also measures the proportion of the total unobserved variability due to unobserved individual heterogeneity. In general, for the specifications with state dependence it is necessarily to assume that the explanatory variables are strictly exogenous, that is, the x_{it} are uncorrelated with u_{it} for all t and s.[18] I also assume that dynamics of decisions are of first order, once x_{it} and c_i are conditioned on. Under these assumptions, the probability of vaccination conditional on the regressors and the unobserved individual effect is:

$$P(d_{it} = 1|d_{i,t-1}, ..., d_{i0}, X_i, c_i) = P(d_{it} = 1|d_{i,t-1}, x_{it}, c_i) = \Phi(x_{it}\beta + \gamma d_{i,t-1} + c_i), \tag{3.12}$$

where $X_i = (x_{it}, x_{i,t-1}, ..., x_{i0})$ and Φ is the standard normal cdf. The second equality follows from the normality assumption of the error term u_{it}.
Given the assumptions, we can write the joint density of the sequence of decisions between period 1 and T, $\mathbf{d}_i = (d_{i1}, ..., d_{iT})$, given (X_i, d_{i0}, c_i) as:

$$f(\mathbf{d}_i|X_i, d_{i0}, c_i; \theta) = \prod_{t=1}^{T} f(d_{it}|d_{i,t-1}, ..., d_{i0}X_i, c_i; \theta) \tag{3.13}$$

$$= \prod_{t=1}^{T} f(d_{it}|d_{i,t-1}, x_{it}, c_i; \theta) \tag{3.14}$$

$$= \prod_{t=1}^{T} \Phi\left[(x_{it}\beta + \gamma d_{i,t-1} + c_i)(2d_{it} - 1)\right] \tag{3.15}$$

The presence of a lagged dependent variable as a regressor together with unobserved heterogeneity raises what has been called the "initial conditions problem", because the first observed decision in the data for individual i, d_{i0}, can be correlated with the unobserved component, c_i, introducing endogeneity problems. Treating c_i as parameters to be estimated, results in inconsistent estimates for β and γ as as N goes to infinity and T is fixed −the incidental parameters problem. To estimate the parameters $\theta = (\beta, \gamma)$, unobserved

[18]The regressors I consider as exogenous are socioeconomic characteristics such as education, marital status, gender, race, and age. I control for chronic illnesses suffered in periods prior to the period in which the first choice d_i0 is observed. I also control for time-varying health measures, like the number of limitations with activities of daily living (ADLS), self-assessed health status, self-reported health changes, and health shocks. These variables are constructed based on survey information prior to the period in which vaccination decision is made, as well as health shocks indicators, which are are constructed using claims in the year prior to the one in which the vaccination decision is made. . Finally, I also assume exogenous whether the individual has a supplemental health insurance and whether it only has Medicare part A (hence, flu vaccination is not free of charge).

heterogeneity c_i must be integrated out. I describe below two solutions proposed in the literature where the estimation of $\theta = (\beta, \gamma)$ is carried out by integrating the unobserved heterogeneity component c_i.

3.4.1 Wooldridge approach

One of the solutions proposed to tackle the initial conditions problem in dynamic, nonlinear panel data models is the conditional maximum likelihood approach proposed by Wooldridge (2005). The procedure consists in finding a density for the sequence of observed choices from period 1 to T, $\mathbf{d}_i$, conditional on the first observed choice, d_{i0}, and all the exogenous variables in all periods, X_i. This can be done finding a density for c_i conditional on d_{i0} and X_i, say $h(c|d_0, X; \delta)$, where δ are the parameters of this density function.
Assuming that $h(c|d_0, X; \delta)$ is the correctly specified density, then the joint density of the sequence of choices $\mathbf{d}_i$, given (X_i, d_{i0}) is:

$$f(\mathbf{d}_i|X_i, d_{i0}; \theta, \delta) =$$
$$= \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T} \Phi\left[(x_{it}\beta + \gamma d_{i,t-1} + c)(2d_{it} - 1)\right] \right\} h(c|d_{i0}, X_i)dc \tag{3.16}$$

We can allow for correlation between the observed regressors and the unobserved individual effect. Following the specification of Mundlak (1978) and Chamberlain (1984), I parameterize the distribution of the unobserved effect as:

$$c_i = \psi + \lambda d_{i0} + \bar{x}_i\alpha + a_i, \tag{3.17}$$
$$a_i \sim N(0, \sigma_a^2), \tag{3.18}$$
$$a_i \quad \text{independent of } d_{i0} \text{ and } \bar{x}_i, \tag{3.19}$$

and $\bar{x}_i$ are the within individual mean (over time) of the time-varying regressors. As indicated in Contoyannis, Jones and Rice (2004), three features of this specification should be noted. First, the specification implies that the identified coefficients for time-invariant regressors are composite effects of the direct effect of these variables on vaccination decision, and the correlation between the time invariant-variable and the unobserved time invariant error term, captured by the relevant elements in β and α. Second, all time dummies must be dropped from $\bar{x}_i$ to avoid perfect collinearity. Finally, the estimate of λ will be of interest to assess the relationship between the initial vaccination decision and the unobserved individual effect.
Under these assumptions the conditional density of c is given by $h(c|d_{i0}, X_i) \sim N(\psi + \lambda d_{i0} + \bar{x}_i\alpha, \sigma_a^2)$ and characterized by the parameters $\delta = (\psi, \lambda, \alpha, \sigma_a^2)$.
For this particular case, the joint distribution of $\mathbf{d}_i$ conditional on observable regressors in

equation (3.16) is:

$$f(\mathbf{d}_i|\bar{x}_i, d_{i0}; \theta, \delta) =$$

$$\int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T} \Phi\left[(x_{it}\beta + \gamma d_{i,t-1} + \psi + \lambda d_{i0} + \bar{x}_i\alpha + a_i)(2d_{it} - 1)\right] \right\} g(a)da = \quad (3.20)$$

$$= \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T} \Phi\left[(x_{it}\beta + \gamma d_{i,t-1} + \psi + \lambda d_{i0} + \bar{x}_i\alpha + a_i)(2d_{it} - 1)\right] \right\} \frac{1}{\sigma_a}\phi(\frac{a}{\sigma_a})da \quad (3.21)$$

where ϕ is the standard normal distribution function.[19] The density in equation (3.21) is the expression of the standard random effects probit model, where the set of regressors is now $W_i = (x_{it}, d_{i,t-1}, d_{i0}, \bar{x}_i)$, and can be estimated as a standard random effects probit model to obtain estimates for $\psi, \beta, \gamma, \lambda, \alpha$ and σ_a^2.[20]

3.4.2 Heckman approach

The approach to the initial conditions problem proposed by Heckman (1981) involves finding a distribution for the first observed choice, $g(d_{i0}|z_{i0}, c_i)$, where z_{i0} is a vector of exogenous instruments. The solution proposed is to specify a linearized approximation to the reduced form equation for the initial value of the latent variable in the following way:

$$d_{i0}^{*} = z_{i0}\pi + \eta_i \quad (3.22)$$

where z_{i0} is a vector of exogenous instruments, which includes pre-sample variables and also the exogenous variables in period 0, x_{i0}. The unobserved component η_i is assumed to be correlated with c_i but not with u_{it}. I assume the following specification:

$$\eta_i = \vartheta c_i + u_{i0}, \quad (3.23)$$

where c_i and u_{i0} are independent. If the initial condition, d_{i0}, is correlated with the unobserved effect, c_i, then $\vartheta \neq 0$, a condition that can be tested. The error term u_{i0} satisfies the same distributional assumptions as u_{it} for $t = 1...T$, that is $u_{i0} \sim N(0, 1)$. Plugging (3.23) into (3.22), the latent variable for the initial period becomes:

[19] Equation (3.21) uses the fact that since $a_i \sim N(0, \sigma_a^2)$ then $\frac{a}{\sigma_a} \sim N(0, 1)$. Then

$$g(a) = \frac{1}{\sigma_a\sqrt{2\pi}} \exp(-\frac{a^2}{2\sigma_a^2}) = \frac{1}{\sigma_a}\left(\frac{1}{\sqrt{2\pi}} \exp(-\frac{(a/\sigma_a)^2}{2})\right) = \frac{1}{\sigma_a}\phi(\frac{a}{\sigma_a})$$

[20] This approach has the advantage that it can be estimated using standard software. In STATA, random effects probit models can be estimated using the `xtprobit` command.

$$d_{i0}^{*} = z_{i0}\pi + \vartheta c_i + u_{i0} \tag{3.24}$$

Then,

$$P(d_{i0} = 1|z_{i0}, c_i) = \Phi(z_{i0}\pi + \vartheta c_i) \tag{3.25}$$

The joint probability of the whole sequence of decisions for individual i, including the initial observation, $(d_{i0}, ..., d_{it})$, given c_i is:

$$L_i \quad = \quad \int \left\{ \Phi(z_{i0}\pi + \vartheta c_i) \times \prod_{t=1}^{T} \Phi\left[(x_{it}\beta + \gamma d_{i,t-1} + c)(2d_{it} - 1)\right] \right\} \frac{1}{\sigma_c}\phi(\frac{c}{\sigma_c})dc \tag{3.26}$$

with $c_i \sim N(0, \sigma_c^2)$. Correlation between the unobserved effect, c_i, and the regressors, X_i, is allowed using, for instance, the Chamberlain-Mundlak method.[21]

3.4.3 Model selection

To assess the statistical fit for different specifications I use the maximized log likelihood. To enable comparison between the results from the Wooldridge and Heckman estimators, the log likelihood of the Wooldridge estimator based on $t \geq 1$ is combined −added− with the log likelihood of a simple probit model estimator for vaccination decision $t = 0$ (Stewart, 2007).[22] The Akaike and Bayesian information Criteria (AIC and BIC) are also reported. These measures capture the trade-off between the model fit −measured by the maximized log likelihood− and the principle of parsimony that favors a simple model. AIC and BIC are calculated as follows:

$$AIC = -2\ln L + 2q, \qquad BIC = -2\ln L + (\ln M)q \tag{3.27}$$

where q represents the number of parameters in each specification and M denotes the number of observations. The difference between the two is that BIC penalizes more the model complexity.

3.5 Results

In this section I present the results for a variety of probit specifications of vaccination models discussed in Section 3.4. Before showing the results for the dynamic models, I estimate static

[21] The integral over c can be evaluated using Gaussian-Hermite quadrature. The program `redprob` (Stewart, 2006) in STATA provides a maximum likelihood estimation of equation (3.26).

[22] The Heckman approach estimates simultaneously the probit model for vaccination decision in $t = 0$ and the dynamic model of decisions for $t \geq 1$. See equation (3.26).

models that do not allow for state dependence. To the extent that observable differences can explain the observed serial persistence in vaccination decisions, very simple static models will be sufficient to explain the participation decision. The results of the static models will provide a benchmark to compare with previous literature that have mainly focused on static models and to assess the role of state dependence in vaccination decisions.

3.5.1 Static Models

Table 3.3 contains the results for models that focus on risk factors and socioeconomic and demographic characteristics as main determinants of flu vaccination. All specifications ignore possible dynamic effects of previous vaccination on current decisions.

Column (1) presents the results of a simple pooled probit model that also ignores possible correlation between decisions in different periods due to time-invariant unobserved heterogeneity. The results indicate that vaccination propensity increases with age for individuals in good health (the baseline category), with a slowdown as age increases (the coefficient of the age squared is negative although not significant).[23]Being in regular health or bad health increase the probability of vaccination for individuals aged 65.

Results from the pooled probit also indicate that preexisting health conditions like cancer, heart disease, diabetes, and chronic lung diseases make individuals more likely to get the vaccine. Suffering respiratory diseases in the previous period increase the probability of current vaccination as well. Individuals may associate the occurrence of this health shock to a higher risk of getting the flu, which in turn increases their likelihood of vaccination, as is found in Ayyagari (2007) and Mullahy (1999). Other health shocks that occurred recently and that individuals may associate to an increase in their risk of complications in case they get the flu, like heart problems and stroke in the last period, do not seem to affect current decisions.

Additionally, the results indicate that physical limitations impose a significant cost in pursuing prevention. Finally, according to this specification, married, white, and more educated individuals are more likely to get the vaccine. Females as well as individuals that have supplementary health insurance have also higher incentives to get the vaccine.

Columns (2) and (3) present the results for random effects probit models that allow for a time-invariant unobserved component in the error term. The correlated random effects (CRE) probit model in column (3) allows for correlation between the explanatory variables and the unobserved heterogeneity following the Mundlak-Camberlain specification. We can see in the last row of column (3) of Table 3.3 that the null hypothesis, stating that the explanatory variables are not correlated with the unobserved time-invariant error term, is rejected. Hence, I will concentrate now in the results of column (3) and compare them with the model in column (1) that ignores this unobserved heterogeneity. The fit with the

[23]Age is normalized relative to the minimum age observed in the sample, which is 65.

Table 3.3: Static Probit Models of Vaccination Decisions.

	Pooled Probit (1)	Random Effects (2)	Correlated Random Effects (3)
Demographic and Socioeconomic Characteristics			
Education	0.185***	0.758***	0.775***
	(0.043)	(0.237)	(0.241)
Female	0.108**	0.285	0.340
	(0.042)	(0.225)	(0.228)
White	0.185***	0.899***	0.885***
	(0.057)	(0.317)	(0.318)
Married	0.231***	0.572***	0.633***
	(0.043)	(0.200)	(0.201)
Supplementary Health Insurance	0.259***	0.310	0.262
	(0.065)	(0.244)	(0.244)
Medicare A Only	-0.025	-0.365	-0.181
	(0.117)	(0.453)	(0.450)
Risk factors: age and subjective health			
Age	0.090***	0.261**	0.349***
	(0.033)	(0.114)	(0.123)
Age^2	-0.004	-0.009	-0.015
	(0.003)	(0.009)	(0.010)
Health Regular	0.528***	0.865**	0.605
	(0.167)	(0.437)	(0.463)
Health Bad	0.346*	0.721	0.503
	(0.204)	(0.523)	(0.558)
Age× Regular	-0.133**	-0.134	-0.069
	(0.053)	(0.137)	(0.145)
Age^2 × Regular	0.010**	0.007	0.002
	(0.004)	(0.011)	(0.012)
Age× Bad	0.011	-0.079	-0.112
	(0.089)	(0.173)	(0.185)
Age^2 × Bad	-0.005	0.002	0.009
	(0.008)	(0.015)	(0.016)
Risk factors: preexisting health conditions			
Cancer ini	0.089*	0.366	0.390
	(0.054)	(0.291)	(0.295)
Heart Conditions ini	0.202***	0.690***	0.669**
	(0.045)	(0.230)	(0.269)
Diabetes ini	0.172***	0.600**	0.524*
	(0.055)	(0.290)	(0.297)
Chronic Lung Disease ini	0.168**	0.744**	0.468
	(0.066)	(0.347)	(0.363)
Stroke ini	-0.019	-0.236	-0.095
	(0.070)	(0.369)	(0.404)
Arthritis Rheumatoid ini	0.104	0.386	0.395
	(0.072)	(0.389)	(0.396)
Risk factors: Health shock previous year			
Heart Disease last year	-0.039	-0.059	-0.031
	(0.062)	(0.145)	(0.151)
Stroke last year	0.171	0.738**	0.747**
	(0.150)	(0.352)	(0.364)
Respiratory Disease last year	0.265***	0.418***	0.274**
	(0.045)	(0.119)	(0.127)
Physical limitations			
Broken hip last year	-0.390*	-1.599**	-1.665**
	(0.232)	(0.645)	(0.702)
Number ADLS	-0.080***	-0.087	-0.006
	(0.019)	(0.058)	(0.067)
Planning horizon			
Health worse than last year	-0.109**	-0.381***	-0.364***
	(0.054)	(0.130)	(0.135)
N	5313	5313	5313
ρ (cross-period correlation)		0.926	0.925
Log-likelihood	-2891.6	-1937.1	-1919.7
AIC	5865.2	3958.1	3951.4
BIC	6134.9	4234.4	4319.8
Wald Statistics for H_0: CRE=0 (p-value)			(0.003)

Note: This table reports the estimated coefficients of probit models. All specifications include year and regions dummies. Age is normalized to be 0 for individuals aged 65, the minimum age observed in the sample. Standard errors are in parenthesis. Pooled probit in column (1) pools all years and assumes iid errors across i and t. The CRE model in column (3) expresses c_i as a linear function of the means of the time-varying regressors. Specifically, $c_i = \psi + \bar{x}_i\alpha + a_i$, where $a_i \sim N(0, \sigma_a^2)$, $\bar{x}_i$ are the within individual mean of the time varying regressors, and a_i is independent of $\bar{x}_i$. The coefficients of the probit model in column 1 are not directly comparable with the RE probits in columns 2 and 3. To make comparisons, multiply the coefficients of the RE models by $(1-\rho)^{1/2}$. In both RE specifications $(1-\rho)^{1/2}$ is approximately 0.27.

CRE probit model largely improves, according to both the AIC and BIC criterion. The model indicates that the estimated unobserved individual heterogeneity, captured by the parameter ρ in equation (3.11), is an important factor for vaccination decisions, accounting for 92.5 percentage of the variability of the error term.

The magnitude of the effect of age and health variables changes when unobserved heterogeneity is allowed. Is worth noting that the coefficients of the pooled probit in column (1) are not directly comparable to those of the random effects probit models in columns (2) and (3) because of the different normalizations of the variance of the error term (Arulampalam, 1998). To make them comparable, one should multiply the coefficients of the RE models by $(1-\rho)^{1/2}$ (0.27 for the CRE probit model of column (3)).

Under the CRE specification in column (3), age is still an important factor determining vaccination of individuals with good health, and the magnitude of the effect is almost the same as in the pooled probit model (the coefficient of age is 0.095 after the adjustment versus 0.090 in the pooled probit). However, the effect of age for individuals in regular and bad health does not seem to be significantly different than for those with good health. The magnitude of the level effect of having regular health is reduced by more than 2/3 in the CRE model (0.165 after the adjustment versus 0.528 in pooled probit) as well as the level effect of being in bad health (0.137 versus 0.346).

Heart disease and diabetes are the only preexisting conditions that significantly affect vaccination decisions when unobserved heterogeneity is introduced. Respiratory shocks and stroke in the last year also increase the probability of getting the vaccine. For all these variables, the magnitude of the effect is lower than in the pooled probit model. Reporting worse health relative to previous period and physical limitations (except for the number of ADLS) are still negatively correlated with the vaccination decision. More educated, white, and married individuals are more likely to pursue prevention, and the magnitude of the effects remains roughly the same than the pooled model.

3.5.2 Dynamic Models

Estimates of the dynamic random-effects probit models of the probability of vaccination using both Wooldridge and Heckman estimators are given in Table 3.4. The two models are also compared to a model that ignores the endogeneity of the initial vaccination decision. The three specifications allow for correlation between the unobserved time-invariant error. Column (1) in Table 3.4 presents the results of the model specification that assumes that the initial condition is exogenous. State dependence takes a predominant role in explaining the time persistence of the vaccination decisions (the coefficient of the lagged dependent variable in absolute value is the largest of all regressors), while the estimated unobserved heterogeneity is irrelevant ($\rho \approx 0$). This pattern changes substantially when the exogeneity assumption is relaxed.

The Wooldridge and Heckman models in columns (2) and (3) account for the "initial condition problem" and in both cases the coefficient of the lagged dependent variable drops dramatically relative to column (1). At the same time, the proportion of the total variance of the errors explained by unobserved heterogeneity raises to 59% according to the Wooldridge approach and to 82% according to the Heckman model. The fit of the model also improves with the Wooldridge or Heckman approach.

The results in Table 3.4 also highlight the importance of accounting for the endogeneity of the initial condition since the correlation between the first observed decision, d_{i0}, and the unobserved component, c_i, is statistically different from zero. In the Wooldridge approach, this correlation is captured by the coefficient associated with the dependent variable in period 0 −the parameter λ in equation (3.17). The estimated value of λ is 2.7, statistically different from zero and approximately 3 times higher than the estimated coefficient of the lagged dependent variable, γ. Under the Heckman specification, the correlation is captured by the parameter ϑ in equation (3.24). The results show that this parameter is approximately 1 and statistically different from zero. Ignoring this correlation I would have attributed a higher effect to state dependence in detriment of unobserved heterogeneity.

According to either AIC and BIC we see that there are not significant differences in the fit to the data that the Wooldridge and the Heckman approaches provide.[24]

Both models differ in the estimated magnitude of some of the coefficients of the explanatory variables. Particularly, when the Heckman specification is used, some characteristics as more education, white, and married, have a higher impact on the probability of vaccination than under the Wooldridge specification. Also age and diabetes have higher impact under the Heckman specification.

3.5.2.1 State dependence and individual unobserved heterogeneity

The dynamic models allow to disentangle two sources of persistence in vaccination decisions: persistence due to unobserved individual heterogeneity and persistence attributed to state dependence. The results from the Wooldridge and the Heckman approaches suggest that both sources are important.

The first row of Table 3.5 compares the estimates of ρ, the proportion of the total unexplained variation that is attributed to the unobserved individual effect, obtained from

[24]I computed the predicted sequence of decisions, to asses the ability of the model to fit the *path* of observed vaccination decisions in three periods using the Person goodness-of-fit statistic computed as

$$GOF = \sum_{s=1}^{J} \frac{(n_s - \hat{n_s})^2}{\hat{n_s}}, \tag{3.28}$$

where n_s is the observed frequency, $\hat{n_s}$ is the predicted frequency, and J is the number of cells for the alternative decision paths (8 possible paths in three years). The Heckman model gives a better fit than the Wooldridge apporach (GoF Static CRE=1311.3, GoF Wooldridge=8915.8, GoF heckman=251.2).

Table 3.4: Dynamic Probit Models of Vaccination Decisions.

	CRE-Initial Conditions Exogenous (1)	Wooldridge Approach (2)	Heckman Approach (3)
Flu shot t-1 (γ)	2.430***	0.926***	0.889***
	(0.052)	(0.144)	(0.156)
Demographic and Socioeconomic Characteristics			
Education	0.102*	0.196*	0.402**
	(0.058)	(0.114)	(0.163)
Female	0.015	-0.046	0.185
	(0.057)	(0.110)	(0.154)
White	0.128*	0.219	0.638***
	(0.072)	(0.140)	(0.223)
Married	0.136**	0.244**	0.446***
	(0.057)	(0.109)	(0.143)
Supplementary Health Insurance	0.247***	0.381**	0.389**
	(0.087)	(0.153)	(0.154)
Medicare A Only	-0.104	-0.130	-0.226
	(0.153)	(0.273)	(0.265)
Risk factors: age and subjective health			
Age	0.005	-0.005	0.139***
	(0.019)	(0.037)	(0.053)
Age^2	0.000	0.001	-0.003
	(0.001)	(0.001)	(0.002)
Health Regular	0.555**	0.590**	0.652**
	(0.230)	(0.288)	(0.294)
Health Bad	0.163	0.170	0.239
	(0.313)	(0.382)	(0.389)
Age × Regular	-0.101**	-0.081	-0.097*
	(0.041)	(0.052)	(0.054)
Age^2 × Regular	0.004**	0.003	0.003*
	(0.002)	(0.002)	(0.002)
Age × Bad	-0.019	0.005	-0.013
	(0.057)	(0.071)	(0.072)
Age^2 × Bad	0.001	0.000	0.001
	(0.002)	(0.003)	(0.003)
Risk factors: preexisting health conditions			
Cancer ini	0.080	0.234	0.221
	(0.072)	(0.143)	(0.184)
Heart Conditions ini	0.076	0.071	0.289
	(0.067)	(0.128)	(0.183)
Diabetes ini	0.097	0.173	0.347*
	(0.073)	(0.142)	(0.206)
Chronic Lung Disease ini	-0.063	-0.267	0.177
	(0.089)	(0.171)	(0.275)
Stroke ini	0.070	0.110	-0.130
	(0.099)	(0.190)	(0.280)
Arthritis Rheumatoid ini	0.191*	0.424**	0.349
	(0.098)	(0.193)	(0.280)
Risk factors: Health shock previous year			
Heart Disease last year	-0.086	-0.047	-0.046
	(0.107)	(0.136)	(0.138)
Stroke last year	0.627**	0.746**	0.770**
	(0.264)	(0.333)	(0.340)
Respiratory Disease last year	0.114	0.210*	0.207*
	(0.087)	(0.112)	(0.114)
Physical limitations			
Broken hip last year	-1.269***	-1.500***	-1.531***
	(0.406)	(0.545)	(0.565)
Number ADLS	0.022	0.003	0.007
	(0.047)	(0.059)	(0.061)
Planning horizon			
Health worse than last year	-0.285***	-0.338***	-0.344***
	(0.095)	(0.120)	(0.122)
N	7,084	7,084	7,084
ρ (cross-period correlation)	0.0000	0.594	0.822
Log-likelihood	-2553.9	-2470.7	-2466.7
AIC	5287.84	5123.45	5125.34
BIC	5905.75	5748.22	5784.44
Wald Statistics for H_0: CRE=0 (p-value)	20.31 (0.062)	76.1 (0.000)	40.01(0.021)
Flu shot ini (λ)		2.748***(0.329)	
ϑ (Correlation between c_i and d_{i0})			1.041***(0.162)

Note: This table reports the estimated coefficients of probit models. All specifications include year and regions dummies. Age is normalized to be 0 for individuals aged 65, the minimum age observed in the sample. Standard errors are in parenthesis. All specifications allows for correlated random effects (CRE), with c_i as a linear function of the means of the time-varying regressors. The log likelihood of the Wooldridge estimator based on $t \geq 1$ is combined with the log likelihood of a simple probit model estimator for $t = 0$.

static and dynamic models. Excluding the dynamic model in column (3) of Table 3.5 that considers the initial condition exogenous, we see that the inclusion of state dependence has important effects on the estimated unobserved heterogeneity, dropping from 0.92 in the static models, to 0.59 and 0.82 in the Wooldridge and Heckman dynamic estimators, respectively. Despite the difference between estimates of ρ in the Wooldridge and Heckman approach is about 20 percentage points, the results from both approaches coincide in that there is still a great proportion of the variability in individual decisions explained by unobserved heterogeneity and not captured by other observable characteristics included in the model.

The effect of state dependence is measured by the coefficient of the lagged dependent variable, γ, and they are shown in the panel B of Table 3.5. All the estimates are positive and statistically significant. In order to assess the effect of the lagged dependent variable on the probability of vaccination, I calculate the partial effect of this variable averaged across all individuals (APEs). Both models, Wooldridge and Heckman, produce almost the same estimate of the APE. Experience with the vaccine in the previous year increases on average between 12 and 13 percentage points the probability of current vaccination.

Table 3.5: State dependence and unobserved heterogeneity

	Static Models		Dynamic Models		
	RE Probit	CRE Probit	CRE Probit (initial cond. exog.)	Wooldridge Approach	Heckman Approach
	(1)	(2)	(3)	(4)	(5)
A. Unobserved heterogeneity					
Cross-period correlation (ρ)	0.926	0.925	0.0000003	0.594	0.822
	(0.008)	(0.008)			
B. State dependence					
Coeff for Flu shot$_{t-1}$ (γ)	-	-	2.430	0.926	0.889
			(0.052)	(0.144)	(0.156)
APE			0.718	0.130	0.114
APE (given $d_{i0}=0$)				0.219	
APE (given $d_{i0}=1$)				0.084	

Notes: Standard deviations in parenthesis. Under the Wooldridge specification, the average partial effect (APE) is estimated by the difference of the counterfactual outcome probabilities taking $d_{i,t-1}$ equal 1 and equal 0, respectively, and evaluating the observed regressors x_{it} at their observed values:

$$N^{-1}\sum_{i=1}^{N}\left((\Phi(x_i\hat{\beta}_a+\hat{\gamma}_a+\hat{\psi}_a+\hat{\lambda}_a d_{i0}+\bar{x}_i\hat{\alpha}_a))-\Phi(x\hat{\beta}_a+\hat{\psi}_a+\hat{\lambda}_a d_{i0}+\bar{x}_i\hat{\alpha}_a)\right) \tag{3.29}$$

where the subscript a denotes that the MLE $\bar{\beta},\hat{\gamma},\hat{\psi},\hat{\lambda},\hat{\alpha}$ are multiplied by $(1+\hat{\sigma}_a^2)^{-1/2}$ and $\hat{\sigma}_a^2$ is the MLE of σ_a^2.
Under Heckman specification the APE is calculated in the same way except that d_{i0} is omitted.

3.5.2.2 Risk factors: age and subjective health status

Table 3.6 reports the estimated average probability of vaccination corresponding to different ages and health status for the Heckman and Wooldridge models. The table also report the results of the static CRE probit model. Figure 3.1 graphically displays these results.

The Wooldrdge and the Heckman approaches give slightly different results. Although vaccination propensities tend to increase with age, the pattern is more clear in the Heckman approach. The later also indicates that there is a slowdown in vaccination propensities at advances ages, for those in "Good" and "Bad" health, as the theoretical model predicts. There are no significant difference between vaccination propensities of individuals with different self-reported health status. This result is similar to the findings in Parente, Salkever, and DaVanzo (2005), who find no effect of self reported health status on vaccination propensities.[25]

Table 3.6 also reports the Person goodness-of-fit statistic,[26] to compare the estimated age-health cells frequencies with the observed data. According to this statistic, the Heckman approach fit the data better than the Wooldridge approach, and gives almost the same fit than the static model.

Table 3.6: Predicted average probabilities of vaccination conditional on health and age.

	Age					
	65	70	75	80	85	90
A. Wooldridge Approach						
Good	0.709	0.672	0.679	0.693	0.714	0.741
Regular	0.773	0.678	0.651	0.656	0.693	0.757
Bad	0.728	0.694	0.705	0.722	0.746	0.775
Goodness of fit	20.7					
B. Heckman Approach						
Good	0.595	0.643	0.716	0.764	0.791	0.800
Regular	0.689	0.624	0.665	0.709	0.754	0.800
Bad	0.631	0.652	0.722	0.772	0.805	0.825
Goodness of fit	14.5					
C. Static CRE Probit						
Good	0.547	0.634	0.717	0.767	0.790	0.790
Regular	0.609	0.620	0.687	0.734	0.763	0.778
Bad	0.577	0.634	0.714	0.765	0.793	0.801
Goodness of fit	13.9					

Note: These probabilities are averages across individuals, and heterogeneity across individuals comes for the part of the unobserved heterogeneity that is correlated with the means of the time varying regressors and with d_{i0} in the Wooldridge approach

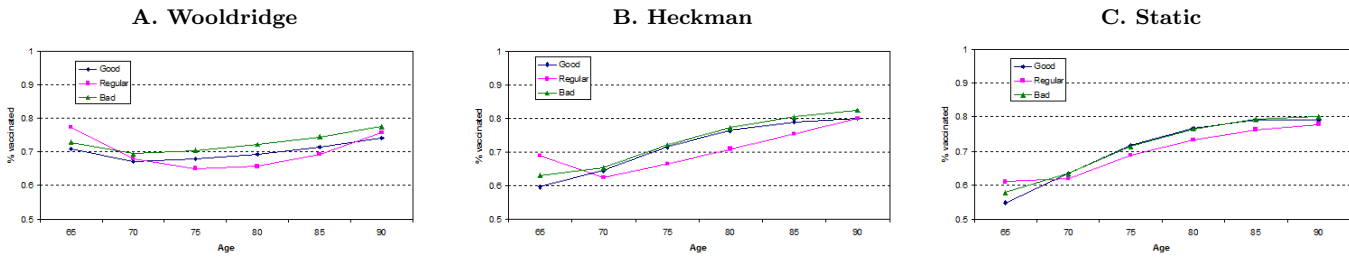
$$N^{-1}\sum_{i=1}^{N}(\Phi(z_i\hat{\beta}_{z,a}+\hat{\beta}_a^{age}\times age+\hat{\beta}_a^{age2}\times age^2+\hat{\beta}_a^{health_j}+\hat{\beta}_a^{age_j}\times age\times health_j \\ +\hat{\beta}_a^{age2_j}\times age^2\times health_j S+\hat{\psi}_a+\hat{\lambda}_a d_{i0}+\bar{x}_i\hat{\alpha}_a) \tag{3.30}$$

where z are all the variables, except age and self-reported health status, $health_j \in \{Regular, Bad\}$. The subscript a denotes that the MLE $\bar{\beta}, \hat{\psi}, \hat{\lambda}$, and $\hat{\alpha}$ are multiplied by $(1+\hat{\sigma}_a^2)^{-1/2}$ and $\hat{\sigma}_a^2$ is the MLE of σ_a^2. Under Heckman specification the APE is calculated in the same way except for the fact that d_{i0} is omitted.

[25] They also use the MCBS dataset, for previous year.

[26] The Pearson goodness-of-fit statistics for the models presented are computed as in equation 3.28, where J is the number of cells (18 for the case of age-health cells).

Figure 3.1: Vaccination rates by age and self-reported health status



3.5.2.3 Risk factors: Preexisting health conditions and health shocks

Results in columns (2) and (3) of Table 3.4 indicate that, once state dependence is introduced, there is in general a reduction of the effect of health conditions −which are related with higher risk of influenza-related complications− on the probability of vaccination relative to the CRE static model (column (3) of Table 3.3). Most of the initial health condions have not statistically significant effect, and the average partial effects in Table 3.7 indicate that the effects are slightly lower effects in dynamic models (except for stroke and arthritis rheumatoid). In all cases the effects of these variables are moderate (of the order of 3 to 5 percentage points).

However, individuals appear to be sensitive to the experience of recent health shocks. I find a positive and significantly different from zero effect of respiratory illnesses and stroke experienced during the period prior to influenza season, and the magnitude of these effects are higher in dynamic models. Those who suffered a respiratory disease in the previous year are between 2.3-2.5 percentage points more likely to take the vaccine in the following year, while those who suffered stroke in the previous year are 7.7-8.4 percentage points more likely.

Table 3.7: Average Partial Effects of Risk Factors.

	Static CRE	Woolddridge	Heckman
Risk factors: preexisting health conditions			
Cancer ini	0.032	0.025	0.026
Heart Conditions ini	0.056	0.008	0.035
Diabetes ini	0.043	0.019	0.041
Chronic Lung Disease ini	0.038	-0.030	0.021
Stroke ini	-0.008	0.012	-0.016
Arthritis Rheumatoid ini	0.032	0.045	0.041
Risk factors: Health shock previous year			
Heart Disease last year	-0.003	-0.005	-0.006
Respiratory Disease last year	0.023	0.023	0.025
Stroke last year	0.059	0.077	0.084

3.5.2.4 Physical limitations and planning horizon

Results in columns (2) and (3) of Table 3.4 also show that there is a significant reduction on vaccination propensity if an individual suffered a broken hip in the previous year or if he reports to have experienced a negative health change since last year. Interpreting the reported change in health status as a proxy for expected longevity (Benitez-Silva and Ni, 2008) the results indicate that individuals experimenting a negative health change are less likely to get the vaccine next period because their planning horizon is shorter. According to the average partial effects in Table 3.8, these two effects are larger in the dynamic models than in the static one. A broken hip in prior year implies a reduction of between 20 and 28 percentage points, according to dynamic models.

Table 3.8: Average Partial Effects of Physical Limitations and Planning Horizon.

	Static CRE	Woolddridge	Heckman
Number ADLS	-0.001	0.000	0.001
Broken hip last year	-0.155	-0.191	-0.213
Health worse than last year	-0.031	-0.038	-0.043

3.5.2.5 Demographic and socioeconomic characteristics

The results suggest that socioeconomic characteristics play an important role in determining individuals' preventive behavior in dynamic models as well. According to column (2) and (3) in Table 3.4, more educated, white, and married individuals are more likely to get annual flu immunization. Table 3.9 presents the average partial effects of these variables. Education has a lower effect in dynamic models than in the static model as one may expect if education is correlated with prior vaccination decisions. There is no a clear pattern between the dynamic and statics models regarding the effects of race and marital status.

Table 3.9: Average Partial Effects of Demographic and Socioeconomic Characteristics.

	Static CRE	Woolddridge	Heckman
Education	0.067	0.022	0.050
Female	0.029	-0.005	0.022
White	0.078	0.024	0.082
Married	0.053	0.027	0.054
Supplementary Health Insurance	0.022	0.043	0.049
Medicare A Only	-0.015	-0.014	-0.028

Finally, individuals with supplementary health insurance coverage are more likely to engage in prevention. Comparing to static models, results in Table 3.9 indicate that these effect are larger in dynamic models. Individuals who have Medicaid part A alone and hence do not have free vaccination coverage, are less likely to get the vaccine, although the coefficients are not statistically different from zero.

3.5.3 Simulated responses

To assess the importance of the dynamic effects generated through state dependence, I analyze what are the long-run gains of a vaccination campaign that is able to increase vaccination coverage in about 20 percentage points in a given period. I perform the following simulation exercise. First, using the Heckman approach estimates (which give a better fit to the data) and taking as given the initial observed vaccination decisions (d_{i0}) and the observed individual characteristics (x_{it}) for $t = 1, 2, 3$), I compute the sequence of predicted vaccination decisions for periods 1 to 3.

Then, I assume that a vaccination campaign implemented in $t = 0$ is able to fully cover all individuals that face a higher risk of influenza related complications (i.e., those with health conditions like cancer, heart problems, stroke, chronic lung disease, and diabetes). In the data, this is equivalent to an increase of about 20 percentage points in the proportion of individuals that take the vaccine in period $t = 0$ (from 66% to 85%). Conditional on the simulated initial conditions and taking as given the observed individual characteristics, I simulate the predicted vaccination decisions in periods 1 to 3. The results are reported in Table 3.10. According to these results, a policy that increases vaccination converage in 20 percentage points in a given period has long lasting effects, increasing vaccination coverage in subsequent periods trhough the state dependence effect. The vaccination coverage is 7 percentage points higher after one year, 4 percentage points higher after two years, and 2 percentage points higher after 3 years. Although positive, the magnitude of the long lasting effect of this policy is moderate, and it may not be enough to sustain high vaccination coverage through time.

Table 3.10: Predicted Vaccination Probabilities using Heckman model

	t=0	t=1	t=2	t=3
Predicted probabilities given observed d_{i0} and x_{it}	65.73	72.05	74.93	80.01
Predicted probabilities given simulated d_{i0} and observed x_{it}	84.47	78.77	79.22	81.99
Difference	18.74	6.72	4.29	1.98

3.6 Conclusion

In this paper I conduct an empirical analysis of the determinants of vaccination decisions and I contrast the results of my estimations with the predictions of a theoretical dynamic model of prevention decisions. The empirical regularity that shows that flu vaccination behavior is highly persistent over time, raises the question of how much of this persistent behavior is explained just by habit persistence or individual unobserved heterogeneity, and how much responds to individuals internalizing the costs and benefits of their vaccination decisions.

My results suggest that the three factors, state dependence, unobserved heterogeneity, and health risks and other individual characteristics, generate persistence. The results also

indicate that ignoring state dependence would result in an overestimation of the effects of health risks and of other individual characteristics (such as education, marital status, and race). I also show that the incentives to vaccinate increase with age and but show a slowdown (according to the Heckman approach) at advanced ages. Conditional on age, there are not significant differences by self-assessed health status.

The implications of these results for public health policy purposes are, at least, twofold. First, the fact that individuals do internalize that certain health risks increases the benefit of vaccination, this channel can be exploited to increase even further vaccination take up. For instance, public campaigns that alert that influenza-related complications are more acute for individuals with certain health conditions, will increase vaccination coverage. Second, any public campaign that induces individuals to get the vaccine for the first time, will have effects in subsequent periods, through the habit persistence channel.

3.7 Appendix

3.7.1 Influenza and Pneumonia: Hospitalization Data

Table 3.11: US Hospital Discharges. Population 65 and older. Average 1997-2006.

Diagnosis		Hospital stays		Hosp. rate (per	Charges per stay $\$^a$ Mean		Δ % 97-06	
		Total	(%)	10,000 hab)	65-84	85+	65-84	85+
Influenza (I)	Principal [b]	12,921	0.10	3.61	9,579	9,698	38%	50%
	Secondary [c]	21,890	0.17	6.12				
Pneumonia (P)	Principal	737,013	5.67	207.1	15,541	14,075	75%	80%
	Secondary	1,337,522	10.29	375.2				
I+P	Principal	749,934	5.77					
	Secondary	1,359,412	10.46					
All hospitalizations		13,001,225	100	3,650.9	18,330	14,835	111%	108%

Notes: [a] Dollars 1997. [b] The principal diagnosis is the condition chiefly responsible for occasioning the admission to the hospital. The principal diagnosis is always the reason for admission.[c] The diagnosis is either the principal diagnosis or an additional condition that coexist at the time of admission. Source: HCUPnet, Healthcare Cost and Utilization Project Agency for Healthcare Research and Quality, Rockville, MD. http://hcupnet.ahrq.gov.

3.7.2 Numerical Example

In this section I present a numerical example of the model presented in Section 3.2, to describe some of its properties. I solve a simple version of a dynamic model of discrete choice and solve the model via dynamic programming. Further assumptions are required to obtain close form solutions to the model, and they are stated here.

Let's call $\mathbf{\Omega}_t$ the vector of state variables of individual i in period t and $\mathbf{\Gamma}_t$ the vector of the *observed* state variables − from the econometrician's point of view − where $\mathbf{\Omega}_t = (h_t, w_t, d_{t-1}, \epsilon_t)$ and $\mathbf{\Gamma}_t = (h_t, w_t, d_{t-1})$, where h_t is health status, w_t are individual characteristics (for simplicity I assume $w_t = age_t$), and d_{t-1} is the vaccination decision in the previous period.[27]

H is the set of possible health levels. There are two possible health states, $H = \{h_g, h_b\}$ and $h_g > h_b$.

Utility. The utility function is a linear function given by:

$$U_t(h_t, d_t) = \begin{cases} h_t + \epsilon_t^0 & \text{if } \ d_t = 0 \\ h_t - c_1 + c_2 d_{t-1} + \epsilon_t^1 & \text{if } \ d_t = 1 \end{cases}$$

[27] I eliminate the subscript i through all the section.

Individuals derive utility from their health level,h_t. They also have to bear some costs, c_1, if they decide to pursue prevention, and c_2 is the cost reduction gained due to past experience with the vaccine. The utility is also affected by idiosyncratic choice-specific preference shifters, ϵ_t^j, which are iid over time, and have a cumulative distribution $G_\epsilon(\epsilon_t)$. The idiosyncratic preference shifters are independent across alternatives and have an extreme value type 1 distribution.

Health production. The probability of contracting the illness, $S(d_t)$, depends on the prevention decision and $S(d_t = 1) < S(d_t = 0)$. Once the health shock,s_t, is realized, I assume that the transition between health states only depends on the current health state. The transition probability from the health state l to health state m, conditional on the realization of the shock s_t is denoted $\pi_{ml}(s_t)$.
The probability of being in health state h_m conditional on current health state h_l and current vaccination decision d_t can be written as:

$$F(h_m|h_l, d_t) = S(d_t) \times \pi_{ml}(s_t = 1) + (1 - S(d_t)) \times \pi_{ml}(s_t = 0), \quad \text{for } m \text{ and } l \in \{g, b\} \tag{3.31}$$

Given previous assumptions, we can show that $F(h_g|h_l, d_t = 1) > F(h_g|h_l, d_t = 0)$ for $l \in \{g, b\}$. Also, the following conditional independence condition also holds: $F(h_{t+1}|h_t, d_t, \epsilon_t) = F(h_{t+1}|h_t, d_t)$.
Determining values for $S(0), S(1), \pi_{gg}(0), \pi_{gg}(1), \pi_{gb}(0)$ and $\pi_{gb}(1)$ we have that the health transition probability function, F, is characterized by 4 values that indicate the probability of being in good health next period conditional on the two possible health states and the current decision.
I further assume that $\pi_{gb}(0) - \pi_{gb}(1) > \pi_{gg}(0) - \pi_{gg}(1)$, which implies that the net effect of vaccination over future health is greater for individuals currently in bad health, i.e., $F(h_g|h_b, 1) - F(h_g|h_b, 0) > F(h_g|h_g, 1) - F(h_g|h_g, 0)$.

Survival probability The probability that an individual will survive to the end of the period is denoted by $p_{su}(h_t, w_t, d_t)$. I assume that the survival probability is zero in period T.

Solving the model. The solution to the model is obtained by backward induction. Two critical assumptions − the conditional independence assumption of the evolution of health states and the distributional assumptions imposed over the unobserved error terms (preference shifters) − allow to obtain a close form analytical solution of the model (Rust, 1994). In particular, the value function of individual i at period t, given the state variables $\mathbf{\Omega}_t$ can written as follows:

$$V_t(\mathbf{\Omega}_t) = \max\{\mathbf{v}_t^0(\mathbf{\Gamma}_t) + \epsilon_t^0;\ \ \mathbf{v}_t^1(\mathbf{\Gamma}_t) + \epsilon_t^1\} \tag{3.32}$$

where $\mathbf{v}_t^j(\mathbf{\Gamma}_t)$ is the "choice specific value function" at period t and depends only on the observable state variables. For a given period t we have $2 \times 2 \times 2$ choice specific value functions ($d_t \in \{0,1\}$, $h_t \in \{h_g, h_b\}$, $d_{t-1} \in \{0,1\}$). These choice specific value function have the following closed form:

$$\begin{gathered}\mathbf{v}_t^j(\mathbf{\Gamma}_t) = u(h_t, d_j, d_{t-1}) + \beta p_{su}(h_t, age_t, d_j) \times \\ \sum_{h_{t+1}} \left\{ \log \left[\sum_{k=0}^{1} \exp\left(\mathbf{v}_{t+1}^k(\mathbf{\Gamma}_{t+1}^j) \right) \right] \right\} P(h_{t+1}|d_j, h_t) \quad \text{for } j \in \{0,1\}\end{gathered} \tag{3.33}$$

An individual i at period t decides to pursue prevention if the following condition holds:

$$\begin{aligned} d_t^*(\mathbf{\Gamma}_t) = 1 \quad &\Leftrightarrow \quad \mathbf{v}_t^0(\mathbf{\Gamma}_t) + \epsilon_t^0 < \mathbf{v}_t^1(\mathbf{\Gamma}_t) + \epsilon_t^1 \\ &\Leftrightarrow \quad \mathbf{v}_t^0(\mathbf{\Gamma}_t) - \mathbf{v}_t^1(\mathbf{\Gamma}_t) < \epsilon_t^1 - \epsilon_t^0 \end{aligned}$$

Since ϵ's are random variables, the optimal decision rule can be expressed in probabilistic terms as:

$$\begin{aligned} Prob\left(d_t^* = 1|\mathbf{\Gamma}_{it}\right) &= Prob\left(\mathbf{v}_t^0(\mathbf{\Gamma}_t) - \mathbf{v}_t^1(\mathbf{\Gamma}_t) < \epsilon_t^1 - \epsilon_t^0\right) \\ &= \int I\left\{\mathbf{v}_t^0(\mathbf{\Gamma}_t) - \mathbf{v}_t^1(\mathbf{\Gamma}_t) < \epsilon_t^1 - \epsilon_t^0\right\} dG_\epsilon(\epsilon_t) \end{aligned}$$

Given that $\{\epsilon_{jt}\}$ are iid type 1 extreme value random variables, the multidimensional integrals in the definition of this conditional choice probability have a close form analytical expression.

$$Prob(d_t = 1|\Gamma_t; \theta) = \frac{\exp\{\mathbf{v}_t^1(\mathbf{\Gamma}_t)\}}{\sum_{j=0}^{1} \exp\{\mathbf{v}_t^j(\mathbf{\Gamma}_t)\}} \tag{3.34}$$

3.7.3 Example

- **Health states:** $H = \{h_g, h_b\} = \{1, 0\}$
- **Vaccination cost:** $c_1 = 0.3$, $c_2 = 0.2$
- **Health transition matrix:** $\mathbf{F}(\mathbf{h_{t+1}}|\mathbf{h_t}, \mathbf{d_t} = \mathbf{0}) = \begin{pmatrix} P_{g,b0} & 1 - P_{g,b0} \\ P_{g,g0} & 1 - P_{g,g0} \end{pmatrix} = \begin{pmatrix} 0.65 & 0.35 \\ 0.9 & 0.10 \end{pmatrix}$

$$\mathbf{F}(\mathbf{h_{t+1}}|\mathbf{h_t}, \mathbf{d_t} = \mathbf{1}) = \begin{pmatrix} P_{g,b1} & 1 - P_{g,b1} \\ P_{g,g1} & 1 - P_{g,g1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.2 \\ 0.95 & 0.05 \end{pmatrix}$$

- **Survival probabilities:** $P_s(h_t, age_t, d_t) = \frac{1}{1+\exp\{x_t\alpha\}}$

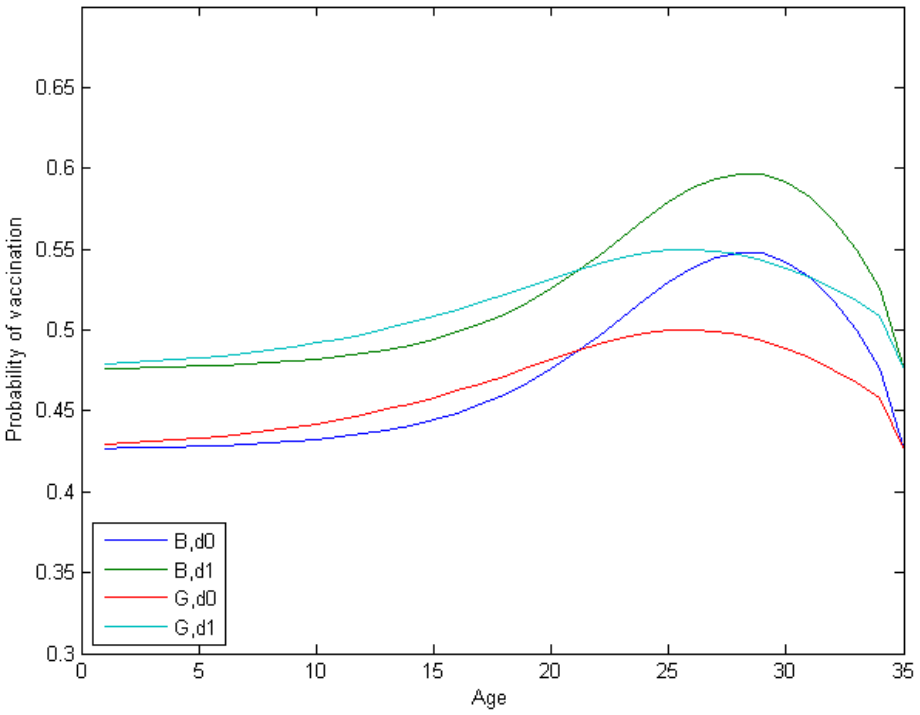
  Where

$$x_t\alpha = \alpha_1 h_t + \alpha_2(1 - h_t) + \alpha_3 age_t + \alpha_4 age_t^2 + \alpha_5 d_t + \alpha_6 d_t(1 - h_t)$$

$$\alpha = (-2.5, -0.7, 0.12, 0.001, -0.1, -0.1)^{28}$$

According to equation (3.34) the conditional probabilities of vaccination, given a health state h, age, and past experience, d_{-1}, are given in figure (3.2)

Figure 3.2: Probabilities of vaccination implied by the model



Note: $(B, d0)$=(Bad Health, $d_{t-1} = 0$), $(B, d1)$=(Bad Health, $d_{t-1} = 1$), $(G, d0)$=(Good Health, $d_{t-1} = 0$), $(G, d1)$=(Good Health, $d_{t-1} = 1$).

[28]Survival probabilities given good health: approximated to the US white female population survival probability profile for ages 65 to 100, period 1999-2001. Source US survival probabilities: National Vital Statistics U.S. Decennial Life Tables for 1999-2001, United States Life Tables. NVSR Volume 57, Number 1. 37 pp. (PHS) 2008-1120.

3.7.3.1 Implications of the model

The implications derived from the model are the following. First, if previous experience with the vaccine reduces the current cost of prevention, then individuals that get the vaccine in one period are more likely to do it again in the following period, compare with those who do not get the vaccine. Second, the probability to get the vaccine increases with age, for a given health status, although there is a slowdown at the end of life. Third, only at advanced ages individuals with worse health status are more likely to get the vaccine relative to those in better health. The gap increases with age and then closes at the end of life.

3.7.4 Variables

VARIABLE	DEFINITION
Flu shot t-1	1 if individual vaccinated last flu season, 0 otherwise
Education	1 if individual completed high school only, or have some college but not diploma, or have bachelor degree or postgraduate degree; 0 otherwise
Female	1 if female, 0 otherwise
White	1 if race white, 0 otherwise
Married	1 if individual is married, 0 otherwise ((widowed, single, divorced or separated))
Insurance	1 if individual have supplemental health insurance, 0 otherwise
Medicare A Only	1 if is only covered by Medicare Part A (not Part B), 0 otherwise
Age	Age in years
Cancer ini	1 if individual has reported having cancer by 2001 (the initial period) or before, 0 otherwise
Heart Disease ini	1 if individual has reported suffering heart disease by 2001 (the initial period) or before, 0 otherwise
Diabetes ini	1 if individual has reported having diabetes by 2001 (the initial period) or before, 0 otherwise
Chronic Lung Disease ini	1 if individual has reported having Chronic Lung Disease by 2001 (the initial period) or before, 0 otherwise
Stroke ini	1 if individual has reported having stroke by 2001 (the initial period) or before, 0 otherwise
Rheumatoid Arthritis ini	if individual has reported having rheumatoid arthritis by 2001 (the initial period) or before, 0 otherwise
Heart Disease, shock last yr	1 if individual that has heart disease experimented a new event related to the illness during the last year, 0 otherwise
Stroke, shock last yr	1 if individual that has stroke experimented a new event related to the illness during the last year, 0 otherwise
Respiratory Illness, shock last yr	1 if individual that has stroke experimented a new event related to the illness during the last year, 0 otherwise
Health change - worse	1 if self reported health compared to previous year is "worse" or "much worse", 0 otherwise ("better", "much better" or "almost the same"
Good Health	1 if self reported health compared to people of same age is "Excellent" or "Very Good", 0 otherwise
Regular Health	1 if self reported health compared to people of same age is "Good", 0 otherwise
Bad Health	1 if self reported health compared to people of same age is "Fair" or "Poor", 0 otherwise
Regions	Dummy variables at the Census Region level: Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, Pacific, or Puerto Rico
Flu shot ini	1 if individual vaccinated in the initial year (2001), 0 otherwise

Bibliography

ACIP (2008): "Prevention and control of influenza. Recommendations of the Advisory Committee on Immunization Practices (ACIP)," Morbidity and Mortality Weekely Report 57.

Almond, D., and J. Currie (2010): "Human Capital Development Before Age Five," NBER Working Papers 15827, National Bureau of Economic Research, Inc.

Angelucci, M., and G. D. Giorgi (2009): "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 99(1), 486–508.

Arulampalam, W. (1998): "A Note on Estimated Coefficients in Random Effects Probit Models," *The Warwick Economics Research Paper Series (TWERPS)*, (520).

Ayyagari, P. (2007): "Do Health Shocks Affect Preventive Behavior?," *SSRN eLibrary*.

Benitez-Silva, H., and H. Ni (2008): "Health status and health dynamics in an empirical model of expected longevity," *Journal of Health Economics*, 27(3), 564–584.

Bhattacharya, J., J. Currie, and S. J. Haider (2006): "Breakfast of Champions?: The School Breakfast Program and the Nutrition of Children and Families," *J. Human Resources*, XLI(3), 445–466.

Black, D., J. Galdo, and J. Smith (2005): "Evaluating the regression discontinuity design using experimental data," *mimeo*, University of Michigan.

Bloom, H. (2009): "Modern Regression Discontinuity Analysis," *mimeo*, MDRC Working Papers on Research Methodology.

Blundell, R., and M. C. Dias (2009): "Alternative Approaches to Evaluation in Empirical Microeconomics.," *Journal of Human Resources*, 44(3), 565 – 640.

Blundell, R., M. C. Dias, C. Meghir, and J. V. Reenen (2004): "Evaluating the Employment Impact of a Mandatory Job Search Program," *Journal of the European Economic Association*, 2(4), 569–606.

Card, D., and L. D. Shore-Sheppard (2004): "Using Discontinuous Eligibility Rules to Identify the Effects of the Federal Medicaid Expansions on Low-Income Children," *The Review of Economics and Statistics*, 86(3), 752–766.

Carneiro, P., and R. Ginja (2009): "Preventing Behavior Problems in Childhood and Adolescence: Evidence From Head Start," *Mimeo*.

Cawley, J., C. Meyerhoefer, and D. Newhouse (2007): "The impact of state physical education requirements on youth physical activity and overweight," *Health Economics*, 16(12), 1287–1301.

Cellini, S. R., F. Ferreira, and J. Rothstein (2010): "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *The Quarterly Journal of Economics*, 125(1), 215–261.

Chamberlain, G. (1984): "Panel Data," Handbook of Econometrics. North- Holland: Amsterdam.

Clark, A. E., and F. Etile (2006): "Don't give up on me baby: Spousal correlation in smoking behaviour," *Journal of Health Economics*, 25(5), 958–978.

Cunningham, P., and A. O'Malley (2009): "Do reimbursement delays discourage Medicaid participation by physicians?," *Health Affairs*, (1), w17–w28.

Currie, J. (2009): "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 47(1), 87–122.

Currie, J., S. Decker, and W. Lin (2008): "Has public health insurance for older children reduced disparities in access to care and health outcomes?," *Journal of Health Economics*, 27(6), 1567 – 1581.

Currie, J., and J. Gruber (1996): "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *The Quarterly Journal of Economics*, 111(2), 431–66.

Cutler, D. M., and J. Gruber (1996): "Does Public Insurance Crowd Out Private Insurance?," *The Quarterly Journal of Economics*, 111(2), 391–430.

Cutler, D. M., and A. Lleras-Muney (2006): "Education and Health: Evaluating Theories and Evidence," NBER Working Papers 12352, National Bureau of Economic Research, Inc.

Decker, S. (2007): "Medicaid physician fees and the quality of medical care of Medicaid patients in the USA," *Review of Economics of the Household*, 5(1), 95–112.

Grieger, L., S. Danziger, and R. Schoeni (2009): "Accurately measuring the trend in poverty in the United States using the Panel Study of Income Dynamics," *Journal of Economic and Social Measurement*, 34(2-3), 105–117.

Gruber, J. (1994): "The Incidence of Mandated Maternity Benefits," *American Economic Review*, 84(3), 622–41.

Gruber, J., and K. Simon (2007): "Crowd-Out Ten Years Later: Have Recent Public Insurance Expansions Crowded Out Private Health Insurance?," Working Paper 12858, National Bureau of Economic Research.

Guryan, J., E. Hurst, and M. Kearney (2008): "Parental education and parental time with children," *Journal of Economic Perspectives*, 22(3), 23–46.

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–09.

Ham, J. C., and L. Shore-Sheppard (2005): "The effect of Medicaid expansions for low-income children on Medicaid participation and private insurance coverage: evidence from the SIPP," *Journal of Public Economics*, 89(1), 57–83.

Harre, N., and A. Coveney (2000): "School-based scalds prevention: reaching children and their families," *Health Education Research*, 15(2), 191–202.

Heckman, J. J. (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process.," in *Structural Analysis of Discrete Data*, ed. by C. F. Manski, and D. Mcfadden. MIT Press.

Imbens, G. W., and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), pp. 467–475.

Imbens, G. W., and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615 – 635, The regression discontinuity design: Theory and applications.

Jacoby, H. G. (2002): "Is There an Intrahousehold "Flypaper Effect"? Evidence From a School Feeding Programme," *Economic Journal*, 112(476), 196–221.

Kahn, E. B., L. T. Ramsey, R. C. Brownson, G. W. Heath, E. H. Howze, K. E. Powell, E. J. Stone, M. W. Rajab, and P. Corso (2002): "The effectiveness of interventions to increase physical activity: A systematic review," *American Journal of Preventive Medicine*, 22(4, Supplement 1), 73 – 107.

Kaiser Commission on Medicaid and the Uninsured (2010): "Medicaid: A Primer," Discussion paper, Kaiser Family Foundation.

Kann, L., N. D. Brener, and D. Allensworth (2001): "Health education: results from the School Health Policies and Programs Study 2000," *Journal of School Health*, 71(7), 266–278.

Kann, L., S. Telljohann, and S. Wooley (2007): "Health Education: Results from the School Health Policies and Programs Study 2006," *Journal of School Health*, 77(8), 408–434.

Kenkel, D. S. (1991): "Health Behavior, Health Knowledge, and Schooling," *Journal of Political Economy*, 99(2), 287–305.

Kenkel, D. S. (2000): "Prevention," in *Handbook of Health Economics*, ed. by A. J. Culyer, and J. P. Newhouse, vol. 1 of *Handbook of Health Economics*, chap. 31, pp. 1675–1720. Elsevier.

Koch, T. (2010): "Using RD Design to Understand Heterogeneity in Health Insurance Crowd-Out," *Working Paper*.

Lalive, R., and A. Cattaneo (2006): "Social Interactions and Schooling Decisions," IZA Discussion Papers 2250, Institute for the Study of Labor (IZA).

Lee, D. S., and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2), 281–355.

Lo Sasso, A. T., and T. C. Buchmueller (2004): "The effect of the state children's health insurance program on health insurance coverage," *Journal of Health Economics*, 23(5), 1059–1082.

Maciosek, M. V., L. I. Solberg, A. B. Coffield, N. M. Edwards, and M. J. Goodman (2006): "Influenza Vaccination: Health Impact and Cost Effectiveness Among Adults Aged 50 to 64 and 65 and Older," *American Journal of Preventive Medicine*, 31(1), 72 – 79.

Maurer, J. (2009): "Who has a clue to preventing the flu? Unravelling supply and demand effects on the take-up of influenza vaccinations," *Journal of Health Economics*, 28(3), 704–717.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142(2), 698 – 714.

McGeary, K. A. (2009): "The Impact of State-Level Nutrition-Education Program Funding on BMI: Evidence from the Behavioral Risk Factor Surveillance System," NBER Working Papers 15001, National Bureau of Economic Research, Inc.

Miguel, E., and M. Kremer (2004): "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72(1), 159–217.

Mullahy, J. (1999): "It'll only hurt a second? Microeconomic determinants of who gets flu shots," *Health Economics*, 8(1), 9–24.

Mundlak, Y. (1978): "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46(1), 69–85.

Nandha, B., and K. Krishnamoorthy (2007): "School-based health education campaign-a potential tool for social mobilization to promote the use of DEC-fortified salt towards elimination of lymphatic filariasis," *Health Education Research*, 22(4), 539–546.

NCQA (2007): "The State of Health Care Quality 2007," Discussion paper, National Committee for Quality Assurance.

——— (2010): "The State of Health Care Quality 2007," Discussion paper, National Committee for Quality Assurance.

Parente, S., D. S. Salkever, and J. DaVanzo (2005): "The role of consumer knowledge of insurance benefits in the demand for preventive health care among the elderly," *Health Economics*, 14(1), 25–38.

Park, C., and C. Kang (2008): "Does education induce healthy lifestyle?," *Journal of Health Economics*, 27(6), 1516 – 1531.

Rust, J. (1994): "Chapter 51: Structural estimation of markov decision processes," vol. 4 of *Handbook of Econometrics*, pp. 3081 – 3143. Elsevier.

Salmon, J., M. L. Booth, P. Phongsavan, N. Murphy, and A. Timperio (2007): "Promoting Physical Activity Participation among Children and Adolescents," *Epidemiologic Reviews*, 29(1), 144–159.

Schmitz, H., and A. Wubker (2010): "What determines influenza vaccination take-up of elderly Europeans?," *Health Economics*, 20.

Shi, X. (2008): "Does an intra-household flypaper effect exist? Evidence from the educational fee reduction reform in rural China," *mimeo*.

Stewart, M. B. (2006): "-redprob- A Stata program for the Heckman estimator of the random effects dynamic probit model," Mimeo.

——— (2007): "The interrelated dynamics of unemployment and low-wage employment," *Journal of Applied Econometrics*, 22(3), 511–531.

STOHR, K. (2003): "The Global Agenda on Influenza Surveillance and Control," *Vaccine*, 21(16), 1744 – 1748, Influenza Vaccine.

TINSLEY, B. J. (2003): *How children learn to be healthy*. Cambridge University Press.

VAN DER KLAAUW, W. (2002): "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 43(4), 1249–1287.

VAN SLUIJS, E. M. F., A. M. MCMINN, AND S. J. GRIFFIN (2007): "Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials," *BMJ*, 335(7622), 703.

WHO (1999): "Improving health through schools: national and international strategies," $http://www.who.int/school_youth_health/media/en/94.pdf$.

——— (2003): "Health and Development Through Physical Activity and Sport," $http://whqlibdoc.who.int/hq/2003/WHO\_NMH\_NPH\_PAH\_03.2.pdf$.

WOOLDRIDGE, J. M. (2005): "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity," *Journal of Applied Econometrics*, 20(1), 39–54.

WU, S. (2003): "Sickness and preventive medical behavior," *Journal of Health Economics*, 22(4), 675–689.

WYATT, T., AND J. NOVAK (2000): "Collaborative partnerships: a critical element in school health programs," *Family & Community Health*, 23(2), 1.