2017

Three Essays in Applied Microeconomics

Xiaohui Guo
Lehigh University

Three Essays in Applied Microeconomics

by

Xiaohui Guo

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Business and Economics

Lehigh University

April, 2017

Certificate of Approval

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

Dissertation Advisor: Chad D. Meyerhoefer

Accepted Date: April 28, 2017

Committee Member:

Chad D. Meyerhoefer

James A. Dearden

Daisy Dai

Arturs Kalnins

Acknowledgments

I have benefited from many people during my doctoral studies. I am very grateful to my advisor, Professor Chad Meyerhoefer and Professor James Dearden for their patient guidance, gracious support, and constant inspiration. More importantly, Chad and Jim have taught me how to be a rigorous researcher and an enthusiastic teacher. It has been a great honor and pleasure working with them as both their student and colleague. I would like to thank Professor Daisy Dai and Professor Arturs Kalnins for being members of my dissertation committee. Their suggestions and insights are very helpful to this thesis. My gratitude also goes to my fellow graduates, especially Lizhong Peng and Mengcen Qian, who had helped me over these past years. Lastly, I am greatly indebted to my family for their love and support.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation consists of three essays. The first essay evaluates restaurant wine pricing. Many restaurants offer extensive lists of wines in bottles and limited selections of wines by the glass. In this essay, I empirically examine wine prices by the bottle at New York City restaurants. The empirical results, which control for both the quality of wine and the restaurant, suggest that restaurants offering wine by the glass tend to set the same bottle prices as the restaurants that do not offer wine by the glass. This equality in price among restaurants is noteworthy because, controlling for quality, restaurants tend to offer a varietal by the glass if and only if they acquire a low-cost brand of the varietal. I then construct a theoretical model where either quantity-based menu pricing or anchoring could explain the bottle pricing practices that I observe in the empirical analysis.

In the second essay, I estimate the effects of prolonged coverage in the State Children's Health Insurance Program (CHIP) during school age on health care utilization and outcomes, using data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999. Using the simulated eligibility as instrument, I identify the causal impacts of CHIP coverage between 1st grade and 8th grade on health care utilization, health status, and academic performance among middle-school children. The results indicate that an additional year coverage in CHIP increases the probability of receiving routine health care by 9 percentage points and the probability of an asthma diagnosis by 7 percentage points. However, I cannot detect any impact of CHIP coverage on health status and academic performance in 8th grade.

In the third essay, I investigate the relationship between human capital accumulation and participation in middle school sports, using data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999. Using the method of instrumental variables, I estimate the causal impact of participation in school-sponsored sports on the academic achievement among middle school children. The results indicate that participation in school sports increases reading test scores, but not test scores in math and science, by 0.27 standard deviations, and that effect is mediated through reduced absenteeism.

Chapter 1: Restaurant Wine Pricing

1.1 Introduction

New York City restaurants Annisa and Robert each offered in July 2015 sauvignon blanc bottles with Wine Spectator ratings of 87 for $56. A 2013 Daniel Reverdy Sancerre -- a sauvignon blanc Annisa lists on its online wine menu -- has an average retail store price of $26, while a 2014 Matua Valley sauvignon blanc -- one Robert posts on its online wine menu -- has an average retail-store price of only $11. This anecdote raises the question of why the restaurant prices for these two identically-rated wines are the same, despite the apparent difference in wholesale prices as evinced by the substantially different average retail-store prices, $26 versus $11. Of course, retail wine store and restaurant prices are both disperse (Jaeger and Storchmann, 2011).[1] We seek to answer this question by focusing on restaurant wine pricing. Restaurants offer food and sell ambiance and experience, which means that identical wine sold by restaurants and retail wine stores could be considered different products. Furthermore, restaurants offer wines by the glass, whereas retail stores mostly do not. We examine the relationship between the features that define restaurants, such as wine-quantity menus and quality, and restaurant bottle pricing.

Continuing with our opening anecdote, Annisa's food and service are more highly rated and the restaurant is more expensive than Robert. Annisa's Zagat ratings are 27 (food), 24 (decor), and 26 (service), while the Robert ratings are 21, 25, and 22. Furthermore, Zagat lists the average cost of a meal at Anissa as $99 and at Robert as $69. Therefore, it is curious that Annisa and Robert charge identical prices for these identically-rated sauvignon blancs, and it is especially interesting

[1] We carefully select our opening anecdote.

that the restaurant and retail-store-average price difference (i.e., markup) is \$30 = \$56 - \$26 at Annisa and a substantially greater, \$45 = \$56 - \$11, at the lower-rated and less-expensive Robert.

In addition to the differences in the food and service ratings and average meal prices of the two restaurants, one potentially important factor that distinguishes Annisa and Robert is that Robert lists its sauvignon blanc on its online by-the-glass menu, while Annisa posts no by-the-glass options on its online wine menu. Hence, Annisa does not promote online its sauvignon blanc by the glass and possibly does not even offer it by the glass.

We examine two factors that affect a restaurant's bottle price of a particular varietal it offers: (a) the quality of the restaurant as measured by the quality of food, decor, and service; and (b) whether the restaurant posts the same varietal by the glass on its online wine menu.

Restaurants vary by the quality and type of food, decor, and service. Hence, the demand for a particular restaurant depends on these factors, and a restaurant's prices therefore depend on these factors as well. In setting prices of food and wine, it is not clear whether a restaurant with better decor, for example, would markup only food prices or both food and wine prices. On one hand, because almost all restaurant goers order food while only some order alcoholic drinks, a restaurant could pass along decor and service costs to all patrons by marking up only food prices. On the other hand, if wine drinkers have more price inelastic demands, then a restaurant with nicer, more-expensive decor may raise wine prices as well.

The relationship between whether a restaurant posts online a varietal by the glass and its bottle price of the varietal is subtle and could involve three issues related to wine pricing: the wholesale prices of the wines offered by the glass, product line options, and anchoring effects.[2] Because

[2] See Adaval and Wyer Jr (2011) for a characterization of consumer behavioral decision processes as a basis for a retailer using a higher-priced product to anchor a consumer's valuation of a lower-priced product. Ariely et al. (2003) demonstrate that initial valuations of familiar products are strongly influenced by even arbitrary anchors.

wine-by-the-glass menus tend to be limited, restaurants have the incentive to offer and promote high-value wines that they secure at low wholesale prices. If they also offer these wines in bottles, then product line and anchoring concerns may drive bottle pricing.

We perform an empirical analysis of wine pricing by New York City restaurants. Recognizing that the quality of a restaurant's food, decor, and service and the types of wines the restaurant offers and promotes by the glass could affect a restaurant's bottle pricing strategy, the goal of our empirical analysis is to examine the relationship between restaurant attributes (food, decor, and service quality and whether restaurants post online by-the-glass wine menus) and measures of restaurant wine-pricing strategies (restaurant wine bottle prices, the average retail store prices of the wines offered by restaurants, and the differences between restaurant and retail wine store bottle prices (i.e., the markup)). We investigate these relationships using data from online by-the-glass and bottle wine menus of New York City restaurants, quality ratings of these restaurants, and average retail wine store prices.

Section 2 presents the data we collected about New York City restaurants. Our empirical method is in Section 3. In Section 4 we describe our results, and in Section 5 we test robustness of our results.

In Section 6 we discuss the empirical relationship that we uncover between bottle pricing and the quality of food, decor, and service. In our empirical analysis, we uncover a relationship between whether a restaurant offers a varietal by the glass and its bottle pricing that warrants a theoretical investigation of the practice. In Section 7 we construct a simple linear city model of oligopolistic competition to investigate pricing and product line issues. In the model with two restaurants and two varietals, each restaurant sells both varietals in bottles. However, for wines by the glass, each restaurant sells only the varietal that it acquires at a low cost. In the context of our model, we demonstrate that either quantity-based menu pricing or anchoring could drive the equilibrium bottle prices of the low-cost wines, which are offered by the glass, and the prices of the high-cost wines, which are not offered by the glass, to be identical. In Section 8 we discuss our results about the relationship between bottle pricing and glass offerings, and we offer suggestions for future research.

1.2 Data

We compile data from New York City (NYC) restaurants that are rated in the 2015 Zagat review. Confining our analysis to NYC allows us to study a set of restaurants in a similar competitive environment with access to a similar set of wholesale suppliers. We further reduce heterogeneity in our sample by limiting the sample to restaurants that serve the following six types of cuisine: American, new American, French, Italian, Mediterranean, and Mexican. These are the most common types of restaurants and they are also the most likely to serve wine. Finally, because we are interested in restaurants that serve wine, we include in our sample only restaurants that Zagat describes as Good for Beer and Wine Only, Good for a Full Bar, Good for Wine, Good for Top Wine, or Good for a Wine Bar.[3] The 2015 edition of Zagat Surveys reviews 2,961 restaurants in NYC. After we subset to restaurants that offer cuisine in the six categories described above and serve wine, our sample contains 741 restaurants, of which 400 post their wine menus online.

From each restaurant's online wine menu, we recorded the producer, vineyard designation (if specified), year, and bottle prices of the least-expensive wines for each of two common white varietals (chardonnay, sauvignon blanc), and two common red varietals (cabernet sauvignon, pinot

[3] We use Zagat scores as a measure of consumer perceptions of restaurant quality. Note that Oliver Gergaud and Verardt (2015) use Zagat ratings to measure changes in consumer perceptions of quality and restaurant investments in response to restaurant ratings by experts.

noir). Not all 400 restaurants list bottle prices for each of these varietals. In particular, 314 restaurants post a price for cabernet sauvignon; 361 post a price for pinot noir; 372 post a chardonnay price; and 357 post a sauvignon blanc price. Counting each bottle price as a single observation, there are 1,404 bottle prices among the 400 restaurants in our dataset that post their wine menus online. Using Wine-Searcher.com, we recorded the average U.S. retail-wine-store prices and the Wine Spectator ratings of these wines reported on the site in July 2015.[4] We use the Wine Spectator ratings as a measure of the quality of each wine.

We also recorded from the online wine menus whether the restaurant offers each of the four varietals by-the-glass. The number of restaurants that post a by-the-glass price on their online wine menus is 202 for cabernet sauvignon; 249 for pinot noir; 281 for chardonnay; and 259 for sauvignon blanc. Across all four varietals there are 991 postings of by-the-glass prices among the 400 restaurants in our dataset with online wine menus.

Zagat Surveys provides the average meal cost at each restaurant.[5] Zagat also offers scores, based on unsolicited consumer reviews, that are normalized to between 0 and 30 for each restaurant's food, decor, and service quality.[6] Zagat maps these scores into the following qualitative characterizations of quality: 0-10 is "poor to fair"; 11-15 is "fair to good"; 16-20 is "good to very good"; 21-25 is "very good to excellent"; and 26-30 is "extraordinary to perfection". Because only 4% of scores are above 25 and below 16, we group these scores into a low score category of less than or equal to 20, and a high score category of above 20.

[4] Wine-Search.com currently lists prices posted by approximately 6,300 U.S. wine stores. The website states that these average prices are calculated using the following methodology: "Average prices are calculated from a 'topped and tailed' data set. We remove the highest and lowest 20% to prevent the average being skewed by pricing errors. When only a small number of prices are available the median is used."
http://www.wine-search.com/average-price.lml

[5] Meal costs are the average consumer-reported prices charged for one dinner including one appetizer, one entree, one drink, and tip.

[6] These 30-point scale ratings were converted to 1-5 star ratings on July 26th, 2016.

1.3 Empirical Method

We use the restaurant bottle price of wine, the average wine store retail price of wine, and the difference between restaurant and retail wine store bottle price (markup) as separate outcomes in our empirical model. Because these price data are available from only restaurants that post their wine menus online, we use a two-part model (TPM; Duan et al. (1983)) to explicitly account for the decision to post the wine menu online and the level of the price outcomes.[7] We specify the first stage as a Probit model of the probability that the restaurant posts a wine menu online:

$$Prob_{ij} = \boldsymbol{\Phi}(\beta_0 + x'_{1ij}\beta_1 + z'_{1ij}\beta_2) \tag{1.1}$$

where i denotes a varietal, j denotes a restaurant, x'_{1ij} is a vector of indicators for a high score for food, service, and decor (i.e., a Zagat score greater than 20), and z'_{1ij} is a vector of control variables for the six different types of cuisine (American, New American, French, Italian, Mexican, and Mediterranean) and 20 different neighborhoods where the restaurants are located.

Because our price outcomes are right-skewed we model the second stage of the TPM using a log-linear OLS regression:

$$\text{Log}(Y_{ij} | Prob_{ij} > 0) = \alpha_0 + x'_{1ij}\alpha_1 + x'_{2ij}\alpha_2 + z'_{1ij}\alpha_3 + z'_{2ij}\alpha_4 + \varepsilon_{ij} \tag{1.2}$$

where Y_{ij} is either the restaurant bottle price, the retail wine store average price, or the markup. The vector x'_{2ij} contains the wine score and an indicator variable that is set to one if the restaurant lists wines by the glass on its online wine menu and set to zero otherwise. The vector z'_{2ij} includes indicators for each varietal (chardonnay, sauvignon blanc, cabernet sauvignon, pinot noir) and for the 20 neighborhoods in NYC.

[7] An alternative would be to use a sample selection model Heckman (1979). However, Manning et al. (1987) demonstrate using a Monte Carlo study that the two-part model outperforms the sample selection model when there is no exclusion restriction to aid identification of the first stage, which is the case here.

The conditional expectation of the three outcomes for the full sample of all NYC restaurant, irrespective of whether they post their wine menus online, is derived as:

$$\mathrm{E}(Y_{ij}|x_{1ij}, x_{2ij}, z_{1ij}, z_{2ij}) = \boldsymbol{\Phi}(\beta_0 + x'_{1ij}\beta_1 + z'_{1ij}\beta_2)\emptyset_s \exp(\alpha_0 + x'_{1ij}\alpha_1 + x'_{2ij}\alpha_2 + z'_{1ij}\alpha_3 + z'_{2ij}\alpha_4) \tag{1.3}$$

where ϕ_s is Duan's nonparametric (Duan, 1983) smearing factor, which is used to retransform the second stage estimates back to the dollar scale. The smearing factor is computed as: $\widehat{\emptyset}_s = \frac{1}{n}\sum_{ij}\exp\left(\hat{\varepsilon}_{ij}\right)$, where is the number of wines posted online, and $\hat{\varepsilon}_{ij}$ is the estimated residual from the second stage of the model.

1.4 Results

First, in Table 1.1 we present summary statistics of the continuous Zagat scores for food, decor, and service; the average restaurant meal cost; the Wine Spectator wine score; and our three price outcomes by whether or not the restaurant posts varietals by-the-glass on its online wine menu.

While the restaurants that post by-the-glass prices on their online wine menus offer similarly-rated wines at similar prices as restaurants that do not post by-the-glass prices, the wines offered by restaurants that post by-the-glass wines have lower wine store retail prices, and as a result, a higher markup.[8] The average difference in the markup between the two groups is \$4.87.

Table 1.2 contains marginal effects of the key regressions calculated for the first part (probability of posting the wine menu online) and second part (price measure conditional on posting the menu online) of the TPM separately as well as the full TPM (unconditional price measure). The standard errors of the marginal effects are clustered at the restaurant level. Only the

[8] The p-value of the t-test over equality of the retail prices between the two groups of restaurants is < 0.00.

decor score is statistically significant in the first stage of the TPM, and indicates that the probability of posting a by-the-glass wine menu online is 11 percentage points higher among restaurants with high decor scores. All of the Zagat scores are precisely estimated in the second stage of the TPM when the outcome is either the bottle markup or the restaurant bottle price. They indicate that higher scores are associated with higher restaurant bottle prices for wine and higher markups among restaurants that post their wine menus online. For example, restaurants in this conditional sample with high decor scores sell bottles for $9.31 more than restaurants with low decor scores, resulting in a $5.43 increase in markup. The markup is also positively associated with higher Wine Spectator scores for the four varietals we consider and is $7.74 per bottle higher when the restaurant posts by-the-glass prices on its online menu.

The total marginal effect estimates that generalize to our full sample of restaurants, irrespective of whether they post their wine menus online, indicate that a high Zagat decor rating is associated with the largest increase in the markup; an increase of nearly $8 per bottle.[9] Most of the difference is due to restaurants with nice decor charging higher bottle prices, although they do also select wines to sell with higher average retail prices. Posting by-the-glass prices is also associated with a $5.89 per bottle increase in the markup. Notably, all of the increase is due to the fact that, for a given Wine spectator rating, restaurants posting by-the-glass prices sell wine with a lower average retail price.

1.5 Generalization of Results

[9] For a dummy variable D, the index function is $x'b + \alpha D$. Let b_1 denote the coefficients of all regressors except for the dummy in the first stage, and b_2 denote the coefficients in the second stage. $\frac{\partial \mathrm{E(y|x)}}{\partial \mathrm{D}} = \emptyset(x'b_1 + \alpha)\emptyset_s \exp(x'b_2 + \alpha) - \emptyset(x'b_1)\emptyset_s \exp(x'b_2)$. For a continuous variable x_1 with a coefficient β_{11} in the first stage and a coefficient β_{21} in the second stage, $\frac{\partial \mathrm{E(y|x)}}{\partial x_1} = \emptyset(x'b_1 + \alpha D)\emptyset_s \beta_{11} \exp(x'b_2 + \alpha D) + \emptyset(x'b_1 + \alpha D)\emptyset_s \beta_{21} \exp(x'b_2 + \alpha D)$.

Our sample contains all the NYC restaurants of six different cuisine types that are reviewed by Zagat. Because maintaining a Zagat rating is costly, the restaurants that are reviewed by Zagat may price wine differently from restaurants that are not reviewed. To investigate this possibility, we collected data on wine prices and restaurant characteristics for all restaurants in the West Village neighborhood of lower Manhattan.[10] We chose this neighborhood to test the generalizability of our results because it contains a large number of restaurants and is relatively easy to define as a market area.

For this sample of restaurants, we use data from Google Maps to control for restaurant cost and quality. Google Maps provides a rating score for total meal cost ranging from one-dollar sign to four-dollar signs, and an overall quality score ranging from one star to five stars. The quality scores are calculated as the average rating submitted by customers to Google. Of the 141 restaurants in the West Village, 91 are reviewed by Zagat and the other 50 are not. Using this sample of restaurants, we estimate a log linear model of markup, restaurant price and retail wine store price with a similar specification to the second part of our TPM. The only difference between this and our earlier model is that the indicators for high Zagat scores for food, decor and service are replaced by indicators for a high overall quality (score ≥ 4), a high meal cost (score ≥ 3), an indicator variable for whether the restaurant was reviewed by Zagat, and exclusion of the dummies for different neighborhoods.

The marginal effect estimates for selected variables are reported in table 1.3. Notably, the indictor variable for whether the restaurant is Zagat-reviewed is imprecisely estimated for all three price measures, although the point estimates suggest a smaller markup among restaurants reviewed

[10] This area is roughly bounded by the Hudson River on the west and Sixth Avenue on the east, extending from West 14th Street south to West Houston Street.

by Zagat. As with Zagat-reviewed restaurants, Google-reviewed restaurants that post by-the-glass wines select wines with lower retail store prices, resulting in a higher markup. However, the increase in markup for restaurants in this particular neighborhood is much lower than in our full sample of Zagat-review restaurants. Overall, these results are consistent with those from our Zagat-reviewed sample, suggesting that our main results are generalizable to all NYC restaurants.

Another concern is that we record the availability of glassed wine options through restaurants' online wine menus. It is possible that restaurants without an online glass program may also promote their glassed wines offline. In that case, online and offline promotions could generate different effectiveness on profitability, including markup (Zhang and Wedel, 2009). We attempt to examine whether there are different effects of by-the-glass option on wine prices and markup between restaurants promoting glass options online and those which may promote glass options offline. We rerun the log linear model of markup, restaurant price and retail wine store price on a subsample of restaurants with an online glassed wine program (i.e., a restaurant offers any of the four varietals by the glass). Notably, the by-the-glass option points to a specific varietal. Of the 338 restaurants in this subsample, they post 989 glassed varietal prices.

The marginal effects of the subsample of restaurants are reported in Table 1.4. The results are very similar to those of our originally full sample. It suggests that the higher markup associated with the by-the-glass option in our baseline model is not driven by the potential difference in sophistication between restaurants do online promotions and those do not.

1.6 Discussion: Bottle Pricing and Restaurant Quality

We believe that possible explanations of the relationships we uncover between the three measures of restaurant quality and the three price measures (the restaurant prices of wine bottles, the retail prices of wines offered by restaurants, and the markup) are straightforward. Considering decor, high scores are associated with higher restaurant bottle prices and higher markups because restaurants could be passing higher costs of decor, which are associated with higher decor scores, at least in part, to wine drinkers. Furthermore, restaurants may optimally set higher wine bottle prices because nicer decor improves the wine experience. The association between service quality and wine prices borders on conventional statistical insignificance and has a weak economic relationship possibly because customers pay for better service through higher tips.[11] High food scores are not associated with bottle prices, which could indicate that restaurants are not passing along the higher cost of high quality food to wine drinkers, possibly because nearly all restaurant goers eat food.

1.7 Theoretical Analysis: Bottle Pricing and Wines By-the-Glass

In a theoretical model with two restaurants, we characterize equilibria that match our empirical results. Specifically, both restaurants set the same prices for the identically-rated bottles of a varietal they offer. They do so in spite of having different costs of the wines - acquiring the wines at different prices. They set the same prices in part because the restaurant that has acquired the varietal at a lower price also offers it by the glass, and the restaurant that acquired the varietal at a higher price, does not. Therefore, empirically and in the context of our theoretical model, despite acquiring the wines at different prices, the glass option at only the restaurant with the lower-cost wine drives the restaurants to set the same bottle prices.

[11] P-values between 0.05 and 0.10 are not considered conventionally low enough to reject the null hypothesis by the statistician Ronald Fisher.

In the model, restaurant goers (consumers) who prefer to drink their favored varietal by the glass drive this pricing equilibrium. Specifically, it is the choices by these light drinkers when their preferred varietal is not offered by the glass that affects the equilibrium prices. When their preferred varietals are not offered by the glass, if they choose glasses of another varietal, then we offer one possible explanation - a rationale that involves a restaurant using its bottle price of a varietal as an anchor for its glass offering of the varietal. Alternatively, if they choose bottles of their preferred varietal when glasses of it are not offered, then we offer another explanation for the equilibrium - a rationale that involves quantity-based menu pricing.

In our simple model, two restaurants, 1 and 2, compete by setting prices for two varietals: merlot (*m*) and pinot noir (*n*). All four wines - both varietals sold by restaurant 1 and both sold by restaurant 2 - have identical ratings. However, to investigate theoretically the relationship that we uncover in our empirical analysis between the prices that restaurants pay for their wines and the prices they set for their customers, the restaurants acquire their wines at different prices. In particular, restaurant 1 has secured a lower-cost *m* and a higher-cost *n* from its supplier, and restaurant 2 has secured a higher-cost m and a lower-cost n from its supplier. We let $\underline{c}$ denote the price a restaurant pays its supplier per bottle of a lower-cost wine, and we let $\overline{c}$ denote the price a restaurant pays its supplier per bottle of a higher-cost wine.

In our empirical analysis we uncover the practice that restaurants tend to offer the wines that they acquire at lower prices, controlling for quality, by the glass. In the spirit of this relationship, we assume in this theory model that each restaurant sells only its lower-cost wine by the glass.[12]

[12] In a model in which the number of wines by-the-glass and in bottles is endogenous, restaurants could offer a limited equilibrium number of wines by-the-glass if either: (1) those who have stronger preferences for limited quantities also have weaker preferences for specific varietals; or (2) due to spoilage from leftover wine in bottles or due to the cost of storage systems, the cost per glass served is increasing in the number of varietals offered.

Specifically, by the glass, restaurant 1 sells only *m* and restaurant 2 sells only *n*. We also assume the quantity of wine per glass is the restaurant standard of 150 ml, which equals 1/5 of a standard 750 ml bottle. The cost for restaurant 1 of selling q_{1bm} bottles and q_{1gm} glasses of *m* and q_{1bn} bottles of *n* is

$$C_1(q_{1bm}, q_{1gm}, q_{1bn}) = \underline{c} \cdot \left(q_{1bm} + \frac{q_{1gm}}{5}\right) + \overline{c} \cdot q_{1bn}.$$

The cost for restaurant 2 of selling q_{2bm} bottles of *m* and q_{2bn} and q_{2gn} bottles of *n* is

$$C_2(q_{2bm}, q_{2gm}, q_{2bn}) = \overline{c} \cdot q_{2bm} + \underline{c}\left(q_{2bn} + \frac{q_{2gm}}{5}\right).$$

Our model is standard in the sense that consumers have preferences for restaurants, quantities of wine, and varietals. However, it is non-standard in that the utility of consuming a glass of a particular varietal at a restaurant may directly depend on the restaurant's bottle price of the varietal. If so, a varietal's bottle price may serve as an anchor for a consumer's willingness to pay for a glass of the varietal.

The restaurants compete in a variant of the standard Hotelling linear model with quantity as well as horizontal differentiation. Our model has a unit mass of consumers, each with a base value, *v*, of dining at one of the two restaurants. The consumers differ in four dimensions. First, a parameter x, which is uniformly distributed on [0, 1], describes consumer tastes for the two restaurants. A consumer located at x on the line spends τx to "travel" to restaurant 1 and spends τ(1-x) to travel to restaurant 2. Second, each consumer is either a heavy drinker or a light drinker. For heavy drinkers, the marginal value of increasing wine consumption from one glass to one bottle is $\theta = \theta_h$; and for light drinkers, the marginal value of increasing wine consumption from one glass to one bottle is $\theta = \theta_l$, where $\theta_h > \theta_l \geq 0$.[13] We let λ denote the proportion of heavy

[13] Our model does not consider restaurant goers in parties of four or more who are light drinkers. In this case, if light drinkers could coordinate on a particular bottle or bottles of wine, then each person could drink only one glass. Incorporating this

drinkers and 1- λ denote the proportion of light drinkers. Third, one-half of the consumers prefer wine *m* to wine *n* and the other half prefer wine *n* to wine *m*. We let $\psi = \overline{\psi} \in (1/2,1)$ denote the value to a consumer from his or her preferred varietal, and $\psi = \underline{\psi} = 1 - \overline{\psi} \in (0,1/2)$ denote the value to a consumer from his or her less-preferred wine. Fourth, consumers differ by their marginal utility of income, which we denote as α. We normalize the marginal utility of income for heavy drinkers at $\alpha_h = 1$, and assume that light drinkers have a marginal utility of income that is no less than one, $\alpha_l > 1$. In our model, heavy drinkers may drink more than light drinkers either because they prefer larger quantities, θ or because they have a smaller marginal opportunity cost of their income, α.

To incorporate the possibility of a price anchor, the utility of consuming a glass of *m* at restaurant 1 includes a component βp_{1bm}, and the utility of consuming a glass of *n* at restaurant 2 includes a component βp_{2bn}.

A consumer located at x with values α, θ, ψ, and β has utility specified in equation (1.4) as follow:

$$U_{x,\theta,\psi} = v + \psi + \theta - \tau x - \alpha p_{1bm} \text{ if one bottle of m at 1;}$$

$$U_{x,\theta,\psi} = v + \psi + \theta - \tau x - \alpha p_{1bn} \text{ if one bottle of n at 1;}$$

$$U_{x,\theta,\psi} = v + \psi + \beta p_{1bm}\theta - \tau x - \alpha p_{1gm} \text{ if one glass of m at 1;}$$

$$U_{x,\theta,\psi} = v + \psi + \theta - \tau(1 - x) - \alpha p_{2bm} \text{ if one bottle of m at 2;}$$

$$U_{x,\theta,\psi} = v + \psi + \theta - \tau(1 - x) - \alpha p_{2bn} \text{ if one bottle of n at 2;}$$

$$U_{x,\theta,\psi} = v + \psi + \beta p_{2bn}\theta - \tau(1 - x) - \alpha p_{2gn} \text{ if one glass of n at 2.} \qquad (1.4)$$

possibility into our model would not affect our qualitative results.

For our analysis, we need to partition the market for restaurant goers into four segments: (1) heavy drinkers who prefer *m*; (2) heavy drinkers who prefer *n*; (3) light drinkers who prefer *m*; and (4) light drinkers who prefer *n*.

To analyze whether menu pricing considerations or anchoring drives the equilibrium prices of bottles to be identical, regardless of the prices paid by restaurants to acquire the wines, we consider two cases about the preferences of light drinkers: (a) they drink only one glass of wine, $\theta_l = 0$; and (b) they drink only their preferred varietal, $\underline{\psi} = 1 - \overline{\psi} = 0$.[14]

Case 1: Anchoring. In this case, the light drinker and heavy drinker markets are segmented in the sense that (a) light drinkers select only one glass and heavy drinkers select bottles and (b) in equilibrium, each segment strictly prefers its selected quantity. With bottle and glass markets segmented in this manner, we show that anchoring can explain the sellers' practice of setting the same equilibrium bottle prices.

To derive the demand function for each restaurant's wine offerings, for each of the four market segments, we need to derive the consumer who is indifferent between dining at restaurants 1 and 2. From the utility function specified in (1.4), these indifferent consumers are the following.[15]

For a light drinker who prefers *m*:

$$\hat{x}_{lm} \equiv \frac{1}{2} + \frac{\left(2\overline{\psi} - 1\right) + \left(p_{2gn} - p_{1gm}\right) + \beta\left(p_{1bm} - p_{2bn}\right)}{2\tau}$$

For a light drinker who prefers *n*:

[14] Note that our goal in evaluating wine pricing theoretically is to provide an explanation of the factor that are related to restaurants tending to set identical bottle prices. In evaluating wine pricing, we demonstrate that an identical-bottle-pricing equilibrium exists for certain parameter values of our model. However, we do not evaluate the robustness of the equilibria. We do not because our goal is to explain our empirical results, where we uncover only that restaurants set the same prices on average. We do not establish that they set the same prices for entire ranges of variables describing consumer preferences.

[15] To simplify the notation in the Case 1 analysis, we set $\alpha_l = 1$.

$$\hat{x}_{ln} \equiv \frac{1}{2} + \frac{\left(1 - 2\overline{\psi}\right) + \left(p_{2gn} - p_{1gm}\right) + \beta(p_{1bm} - p_{2bn})}{2\tau}$$

For a heavy drinker who prefers *m*:

$$\hat{x}_{hm} \equiv \frac{1}{2} + \frac{(p_{2bm} - p_{1bm})}{2\tau}$$

For a heavy drinker who prefers *n*:

$$\hat{x}_{hn} \equiv \frac{1}{2} + \frac{(p_{2bn} - p_{1bn})}{2\tau}$$

With these cutoffs (i.e., demand functions) in place, restaurant 1's profit function is

$$\pi_1 = \left(p_{1gm} - \frac{\underline{c}}{4}\right)(\hat{x}_{lm} + \hat{x}_{ln}) + \left(p_{1bm} - \underline{c}\right)\hat{x}_{hm} + (p_{1bn} - \overline{c})\hat{x}_{hn}$$

and restaurant 2's profit function is

$$\pi_2 = \left(p_{2gn} - \frac{\underline{c}}{4}\right)(2 - \hat{x}_{lm} - \hat{x}_{ln}) + (p_{2bm} - \overline{c})(1 - \hat{x}_{hm}) + \left(p_{2bn} - \underline{c}\right)(1 - \hat{x}_{hn})$$

To simplify our analysis, we characterize a Nash equilibrium for a case in which a heavy drinker's value of a bottle is "large" in the sense that in equilibrium, a heavy drinker strictly prefers consuming a bottle to consuming a glass. With regard to light drinkers, by assumption $\theta_l = 0$ and $\underline{\psi} > 0$, they do not consume bottles. As a result, we can solve for the equilibrium glass prices separately from the equilibrium bottle prices.

If a heavy drinker's marginal value of a bottle, θ_h, is sufficiently large,

$$\theta_h > \max\{\frac{7}{15}\underline{c} + \frac{1}{3}\overline{c} + \frac{2}{3}\beta\frac{1-\lambda}{\lambda}, \left(1 - 2\overline{\psi}\right) + \frac{2}{15}\underline{c} + \frac{2}{3}\overline{c} + \frac{4}{3}\beta\frac{1-\lambda}{\lambda}\}$$

then in a Nash equilibrium, the prices of the restaurants' low-cost bottles are

$$p_{1bm}^* = p_{2bn}^* = \tau + \frac{1}{3}(2\underline{c} + \overline{c} + 2\beta\frac{1-\lambda}{\lambda})$$

the prices of the restaurants' high-cost bottles are

$$p^*_{1bn} = p^*_{2bm} = \tau + \frac{1}{3}\left(\underline{c} + 2\overline{c} + 4\beta\frac{1-\lambda}{\lambda}\right)$$

and the prices of the restaurants' glasses are

$$p^*_{1gm} = p^*_{2gn} = \tau + \frac{\underline{c}}{5}$$

We are concerned with comparing the prices of low-cost bottles (which the restaurants also offer by the glass) and high-cost bottles (which they do not offer by the glass). At each restaurant the equilibrium price of the high-cost wine less the equilibrium price of the low-cost wine is:

$$p^*_{1bm} - p^*_{1bn} = p^*_{2bn} - p^*_{2bm} = \frac{1}{3}(\overline{c} - \underline{c}) - \frac{2}{3}\beta\tau\frac{1-\lambda}{\lambda}$$

Without an effect of a restaurant's bottle price on the utility of consuming a glass (i.e., an anchoring effect), $\beta = 0$, the restaurants' equilibrium bottle prices of the high-cost wines are greater than the equilibrium bottle prices of the low-cost wines. However, with an anchoring effect, $\beta > 0$, it is possible that the restaurants set identical bottle prices for their low-cost and high-cost wines. Specifically, if the magnitude of the anchoring effect on the bottle prices of low-cost wines, $\frac{2}{3}\beta\tau\frac{1-\lambda}{\lambda}$, equals the 1/3 the difference in the cost of high-cost and low-cost wines, $\frac{1}{3}(\overline{c} - \underline{c})$, then $p^*_{1bm} = p^*_{1bn} = p^*_{2bn} = p^*_{2bm}$.

Case 2: Quantity-Based Menu Pricing. In this case, light drinkers select only their preferred varietals. Thus, for the high-cost varietal, which a restaurant offers in bottles only, it prices the varietal to attract not only heavy drinkers who prefer this high-cost varietal, but also light drinkers who also prefer it. However, for its low-cost varietal, which it offers in bottles and by the glass, it prices bottles of this varietal to attract only heavy drinkers who prefer this low-cost varietal, and it prices glasses to attract only light drinkers who prefer it. As we demonstrate, this case can create a scenario in which each restaurant sets the same equilibrium price for its low- and high-cost varietals.

To derive the demand function for each restaurant's wine offerings, for each of the four market segments, we need to derive the consumer who is indifferent between dining at restaurants 1 and 2. From the utility function specified in (1.4), these indifferent consumers are the following. For a light drinker who prefers *m*:

$$\hat{x}_{lm} \equiv \frac{1}{2} + \frac{-\theta_l + \alpha_l\left(p_{2gm} - p_{1gm}\right)}{2\tau}$$

For a light drinker who prefers *n*:

$$\hat{x}_{ln} \equiv \frac{1}{2} + \frac{\theta_l + \alpha_l\left(p_{2gn} - p_{1gn}\right)}{2\tau}$$

For a heavy drinker who prefers *m*:

$$\hat{x}_{hm} \equiv \frac{1}{2} + \frac{(p_{2bm} - p_{1bm})}{2\tau}$$

For a heavy drinker who prefers *n*:

$$\hat{x}_{hn} \equiv \frac{1}{2} + \frac{(p_{2bn} - p_{1bn})}{2\tau}$$

With these cutoffs (i.e., demand functions) in place, restaurant 1's profit function is

$$\pi_1 = \left(p_{1gm} - \frac{\underline{c}}{4}\right)\hat{x}_{lm} + \left(p_{1bm} - \underline{c}\right)\hat{x}_{hm} + (p_{1bn} - \overline{c})(\hat{x}_{ln} + \hat{x}_{hn})$$

and restaurant 2's profit function is

$$\pi_2 = \left(p_{2gn} - \frac{\underline{c}}{4}\right)(1 - \hat{x}_{ln}) + (p_{2bm} - \overline{c})(2 - \hat{x}_{lm} - \hat{x}_{hm}) + \left(p_{2bn} - \underline{c}\right)(1 - \hat{x}_{hn})$$

As in Case 1, we characterize a Nash equilibrium for parameter values in which a heavy drinker strictly prefers consuming a bottle to consuming a glass (θ_h is sufficiently large). In addition, a light drinker strictly prefers consuming a bottle of her preferred varietal to a glass of her less-preferred varietal (θ_l is sufficiently large and $\underline{\psi} = 0$). As a result, we can solve for the Hotelling equilibrium bottle prices separately from solving for the Hotelling equilibrium glass prices.

For certain parameter values of our model, a Nash equilibrium exists in which the firms set identical bottle prices for each of the varietals. Specifically, for any $\theta_h > \theta_l$, $\tau > \overline{c} - \underline{c}$, and $\alpha \geq 1$, if

$$\theta_l = \frac{7}{10}\alpha_l\underline{c} + \frac{1}{2}\alpha_l\tau + \tau$$

$$\lambda = 1 - \frac{20(\overline{c} - \underline{c})}{20(1-\alpha_l)\left(\overline{c} - \underline{c}\right) + 25\alpha_l\tau - 40\tau + \alpha_l\underline{c}}$$

and in equilibrium $\hat{x}_{lm}, \hat{x}_{ln} \in (0,1)$ in particular, then a Nash equilibrium exists in which the restaurants set prices:

$$p^*_{1gm} = p^*_{1gn} = p^*_{2gm} = p^*_{2gn} = \underline{c} + \tau$$

and

$$p^*_{1gm} = p^*_{2gn} = \frac{\underline{c} + \tau}{4}$$

Summary of Cases and 1 and 2. We characterized equilibria for Cases 1 and 2 in which the two restaurants set identical bottle prices for the two varietals, m and n. In these two cases, each restaurant sells only its low-cost wine (*m* for restaurant 1 and *n* for restaurant 2) by the glass.

In Case 1, if a restaurant does not offer a light drinker's preferred varietal by the glass, then he or she orders a glass of the other varietal. Therefore, if light drinkers have a strong preference to drink only a glass of wine and a weak preference for varietals, then we demonstrate that anchoring can cause the empirical pricing practice that we uncover. In Case 2, if a restaurant does not offer a light drinker's preferred varietal by the glass, then he or she orders a bottle of this preferred varietal. Therefore, if light drinkers have a strong preference for varietals and a weak preference to drink only a glass of wine, then we demonstrate that a price-quantity menu (i.e., a glass price and a bottle price) for the low-cost varietal and only a bottle price for the high-cost varietal can cause the empirical pricing practice that we uncover.

1.8 Discussion: Bottle Pricing and Wines By-the-Glass

The association we uncover between whether a restaurant offers a varietal by the glass and the bottle pricing and markup of the varietal we believe is most interesting. It points to a remaining question about restaurant wine offering and pricing strategy. Which rationale - price anchoring or quantity-based menu pricing - drives the restaurants to set identical prices for lower-cost wines offered by the glass and higher-cost wines offered only in bottles? It may be that restaurants purchase greater quantities of wines with lower retail store prices, but instead of passing those savings along to consumers in the form of lower bottle prices, they maintain higher prices in order to make these wines by the glass look more appealing to their customers. Alternatively, or as a complementary pricing strategy, restaurants may set prices of the higher-cost wines they serve only in bottles at a level to encourage light drinkers to purchase bottles. The former strategy is consistent with anchoring and the latter with quantity-based menu pricing.

Which of these two strategies explains why restaurants have higher markups for wines offered by the glass, in the context of our theoretical model, depends on the preferences of light drinkers. In Case 1 of our theoretical model, light drinkers consume only wines by the glass; and in Case 2, they consume only their preferred varietal. Therefore, the Case-1 anchoring rationale for the equilibrium pricing requires that light drinkers have a strong preference to drink only one glass of wine, and the Case-2 menu pricing rationale for the equilibrium pricing requires that light drinkers have a strong preference to drink their preferred varietal, regardless of the quantity.

An interesting issue for future research is about whether light wine drinkers have stronger preferences for quantity or varietals. As we demonstrated, determining these preferences would be critical in determining whether a standard menu pricing argument explains our results or a behavioral anchoring argument does so. Although we cannot distinguish these mechanisms in our empirical analysis, our own view is that anchoring is the more likely explanation. It is because we believe that light drinkers probably have less knowledge of wine and weaker preferences for specific varietals, thereby causing them to favor quantity. Identifying these causal pathways is a fruitful area of research for future studies that make use of data on restaurant wine sales and consumer attributes.

1.9 Tables and Figures

Table 1.1 Summary Statistics

	Restaurants without By-the-Glass Varietal Posted				
VARIABLES	Mean	Std. Dev.	Min.	Max.	N
Food Score	22.591	2.362	15	28	413
Decor Score	19.288	2.780	12	27	413
Service Score	20.998	2.270	15	27	413
Average Meal Price	55.801	15.565	23	128	413
Wine Score	87.004	3.160	60	93	413
Restaurant Wine Price	58.523	26.380	17	180	413
Retail Wine Price	24.666	13.414	6	119	413
Markup	33.857	18.105	2	105	413

	Restaurants with By-the-Glass Varietal Posted				
VARIABLES	Mean	Std. Dev.	Min.	Max.	N
Food Score	21.972	2.542	13	29	991
Decor Score	20.030	3.085	13	28	991
Service Score	20.660	2.582	14	28	991
Average Meal Price	56.677	21.475	17	167	991
Wine Score	86.600	2.626	75	93	991
Restaurant Wine Price	57.298	19.016	20	160	991
Retail Wine Price	18.569	8.881	5	96	991
Markup	38.728	14.600	2	116	991

Table 1.2: Marginal Effects from Two Part Model - All Wines

	(1) Markup	(2) Restaurant Price	(3) Retail Price
Probability of Wine Menus Online			
High Food Score	0.005	0.005	0.005
	(0.033)	(0.033)	(0.033)
High Decor Score	0.107***	0.107***	0.107***
	(0.026)	(0.026)	(0.026)
High Service Score	0.034	0.034	0.034
	(0.027)	(0.027)	(0.027)
Price Measure (Conditional on Posting Online)			
High Food Score	2.261***	4.133***	2.725***
	(1.178)	(1.145)	(0.518)
High Decor Score	5.429***	9.309***	3.457***
	(1.339)	(1.197)	(0.517)
High Service Score	1.579*	1.818**	0.284
	(0.989)	(1.072)	(0.518)
Wine Score	0.399***	1.851***	1.330***
	(0.154)	(0.185)	(0.122)
By-the-Glass	7.736***	1.003	-5.668***
	(1.166)	(1.208)	(0.587)
Unconditional Price Measure			
High Food Score	1.954	3.460	2.176***
	(1.755)	(2.171)	(0.812)
High Decor Score	7.969***	13.035***	4.737***
	(1.868)	(2.117)	(0.775)
High Service Score	2.429*	3.341*	0.923
	(1.452)	(1.919)	(0.762)
Wine Score	0.297**	1.401***	1.007***
	(0.118)	(0.165)	(0.106)
By-the-Glass	5.886***	0.739	-4.363***
	(1.146)	(1.254)	(0.567)
N	1,959	1,959	1,959

Standard errors in parentheses clustered at the restaurant level

*** p<0.01, ** p<0.05, * p<0.1

Table 1.3: Test of Selectin into Zagat Review Sample

	(1) Markup	(2) Restaurant Price	(3) Retail Price
Log-Linear Regression			
High Meal Cost	0.269***	0.244***	0.224***
	(0.095)	(0.052)	(0.066)
High Overall Quality	0.025	0.102	0.252**
	(0.107)	(0.076)	(0.097)
Wine Score	0.008***	0.011***	0.014**
	(0.002)	(0.003)	(0.006)
By-the-Glass	0.319***	0.034	-0.233***
	(0.112)	(0.065)	(0.077)
Zagat Review	-0.099	-0.029	0.068
	(0.097)	(0.072)	(0.088)
N	235	235	235

Standard errors in parentheses clustered at the restaurant level

*** p<0.01, ** p<0.05, * p<0.1

Table 1.4: Marginal Effects of Restaurants with a Glass Program

	(1) Markup	(2) Restaurant Price	(3) Retail Price
Log-Linear Regression			
High Food Score	-1.107	0.862	2.133***
	(1.091)	(1.220)	(0.545)
High Decor Score	7.141***	10.243***	3.422***
	(1.093)	(1.162)	(0.554)
High Service Score	1.130	1.862	0.305
	(1.061)	(1.190)	(0.577)
Wine Score	0.877***	2.360***	1.380***
	(0.162)	(0.209)	(0.110)
By-the-Glass	6.051***	-1.049	-6.242***
	(1.188)	(1.443)	(0.705)
N	1,382	1,382	1,382

Standard errors in parentheses clustered at the restaurant level

*** p<0.01, ** p<0.05, * p<0.1

Chapter 2: The effect of coverage through the State Children's Health Insurance Program on the health and academic performance of middle school children

2.1 Introduction

The State Children's Health Insurance Program (CHIP), established in 1997 as a joint federal-state program, aims to cover uninsured children in near-poor and moderate income families that do not meet the income and categorical eligibility requirements imposed by the Medicaid program. The benefits of providing low-cost health insurance to low-income children is of considerable policy interest, as evidenced by the subsequent expansion of CHIP eligibility (Rosenbach et al., 2007). This expansion increased the likelihood of coverage in CHIP, which may in have facilitated better access to medical care. Better access to care could have benefited children in various ways, for example, by raising their level of health and productivity, or liberating financial resources to fund education, which is a key determinant of human capital (Grossman, 1972).

The economics literature contains several recent studies that evaluate the long-run effects of Medicaid and CHIP eligibility on health, educational attainment, and labor market outcomes (Brown et al., 2015; Cohodes et al., 2016; Meyer and Wherry, 2016; Bourdreaux et al., 2016; Thompson, 2017). These papers conclude that expanded eligibility for Medicaid and CHIP results in more health service utilization, better health status and educational attainment, and higher wages in adulthood. However, the take-up rates of the public welfare programs, including CHIP, are typically well below 100% (Culter and Gruber, 1996; Sasso and Buchmueller, 2004; Dillender, 2017). Therefore, evaluation of the direct impacts of actual program coverage, rather than program eligibility, is important. There are few studies that have investigated the impacts of CHIP coverage on health and educational outcomes for middle school children,[16] and none that we are aware of that use nationally representative data.[17]

Human capital accumulation in adulthood builds upon the school-age experience, and there is reason to believe that CHIP coverage during school-age helps narrow the gap in later life health and educational attainment of between children who grow up in the low and high income households (Case et al., 2002; Currie, 2009). In addition, local communities are more likely to provide routine health care to preschool-aged children without health insurance coverage than school-aged children (Slifkin et al., 2002), leaving more room for CHIP to better the condition of the school-aged children without coverage, and/or those covered by poor quality private plans.

Our paper contributes to the literature by providing the first causal estimates of the effects of the school-age CHIP coverage duration on health care utilization, health status, and academic performance for children in middle school. We analyze data from the nationally representative Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K, 1998). We limit the sample to children with household income between 100% and 300% of the federal poverty line (FPL) because this is the group of children targeted by CHIP (Sasso and Buchmueller, 2004). To address the endogeneity of coverage in CHIP, we use simulated eligibility as the instrument (Currie and Gruber, 1996a, 1996b; Cutler and Gruber, 1996), exploiting variation of CHIP income eligibility across states and over time. Results from our IV estimation indicate that the effects of

[16] Thompson (2017) finds that childhood coverage of both Medicaid and CHIP leads to better adult health outcomes using NLSY-79 data which surveyed children of age 14-21 as of 1979. Our study differs from Thompson (2017) by focusing on CHIP alone and evaluating childhood outcomes.

[17] Cullen et al. (2005) report the medium impact of duration of CHIP coverage on health care utilization and outcomes among children up to 3rd grade from reduced form estimates, using data from ECLS-K, 1998. Low power of the instrument they use in IV estimation due to the use of former waves of the ECLS-K, 1998 hampers causal inference.

CHIP coverage over a multi-year period on accessing consistent routine care and on having diagnosed with asthma are positive and statistically significant. However, we are unable to detect any effect of CHIP coverage on parent-evaluated child health status, obesity, and changes in academic test scores over time.

The rest of the paper is organized as follows: we review the literature of CHIP benefits in Section 2. We describe the data in Section 3 and our empirical approach in Section 4. Section 5 contains our empirical results. In Section 6, we discuss the results and conclude.

2.2 Literature Review

States have the option to implement CHIP as an expansion to the current Medicaid program, as a separate program, or as a combined program.[18] We are unaware of any paper that identifies differential impacts of separate CHIPs and Medicaid-extended CHIPs.[19] Below we discuss the existing literature focusing on the effects of Medicaid and CHIP collectively.

2.2.1 Coverage

A number of research studies have established the ways by which expansions in income eligibility result in a rise in coverage in Medicaid and CHIP, either through the gain of public insurance by the previously uninsured or through the movement of individuals from private plans to public insurance (crowd-out) (Cutler and Gruber, 1996; Sasso and Buchmueller, 2004; Seldon et al., 2004; Cullen et al., 2005; Ham and Shore-Sheppard, 2005; Lee et al., 2008; Gruber and Simon, 2008). These papers consistently document a positive take-up rate under the expansions of

[18] Separate CHIPs have a higher federal matching fund rate than ones that expand the current Medicaid program. In addition, a separate CHIP has more flexibility in designing its benefit, premium, cost sharing, and managed care. For example, the Early Periodic Screening Diagnosis and Treatment Program is required under Medicaid, but not CHIP.

[19] For example, Sasso and Buchmueller (2004) find similar increases in the coverage from expansions in income eligibility from Medicaid based CHIP and from separate CHIP.

Medicaid and CHIP. However, evidence on the extent of the crowd-out effect is mixed, ranging from 0% to 60%.[20] For example, Ham and Shore-Sheppard (2005) find no significant decline in private insurance coverage associated with Medicaid expansions. In contrast, Sasso and Buchmueller (2004) report a joint crowd-out rate of Medicaid and CHIP of just under 50%, which is in line with the estimated Medicaid crowd-out rate by Cutler and Gruber (1996).

Although the goal of CHIP expansions is to reduce the rate of uninsured children while limiting crowd-out,[21] CHIP reaches near poor and moderate income households where over half of children are already covered by private insurance plans (Cullen et al., 2005). The previous literature focuses on families with household income between 100-200% of FPL in the 1990s (e.g., Cutler and Gruber, 1996). We provide additional evidence on the CHIP take-up and crowd-out rates among children with a higher income range of 100-300% of FPL following the CHIP expansion in the 2000s.[22]

2.2.2 Health Care Utilization

There is a large body of literature that examines the effects of Medicaid and CHIP expansions in eligibility on health care utilization among low income children (Currie and Gruber, 1996a; Banthin and Seldon, 2003; Davidoff et al., 2005; Currie et al., 2008; Meyer and Wherry 2016; Bourdreaux et al., 2016; Thompson 2017). Most of these studies provide evidence that expansions in eligibility results in better access to health care, including routine medical care and dental care among eligible children.[23] The evidence of the effect of program coverage on health care

[20] The upper limit of the range of crowd-out ratios is identified in Gruber and Simon (2008).

[21] Gruber and Simon (2008) find that anti-crowd-out measures, such as the frozen window of transition from private insurance to CHIP, generate counter effects by lowering take-up rates more than crowd-out rates.

[22] Most of the states cover children of household income up to 300% of the FPL under CHIP. One exception is New Jersey, which institutes the income eligibility of CHIP at 350% of the FPL during our study period.

[23] These studies use either IV or Difference-in-Difference (DID) to derive causal inference. While other studies identify positive effect of Medicaid/CHIP eligibility on access to health care, Banthin and Seldon (2003) shows no significant improvements in access of use of health services among CHIP eligible children. Neither there is no improvement in financial burden among CHIP

utilization is consistently positive (e.g., Fox et al., 2003; Dick et al., 2004;), though almost all are state-specific analysis.

One exception is Seldon and Hudson (2006), who use the 1996-2002 Medical Expenditure Panel Survey (MEPS) data to study the nationwide effects of coverage through Medicaid and CHIP for children under the age of 18. Using Instrumental Variable (IV) estimation,[24] they identify a positive effect of current coverage on current access to health care. However, these estimates do not reflect dynamics of health care utilization. Our analysis provides estimates of the duration of CHIP coverage on health care access across several years.

2.2.3 Health Status and Educational Attainment

The strand of literature that examines the impact of Medicaid and CHIP on health outcomes generates mixed evidence depending on the outcome measures analyzed (Currie and Gruber 1996a, 1996b; Racine et al., 2001; Lykens and Jargowsky, 2002; Cullen et al., 2005; Dafny and Gruber, 2005; Currie et al., 2008; Levine and Schanzenbach, 2009; Howell et al., 2010; Cohodes et al., 2016; Meyer and Wherry; 2016; Thompson, 2017). For example, there is consistent evidence of reductions in child mortality associated with the Medicaid and CHIP expansions (Currie and Gruber 1996a, 1996b; Howell et al., 2010). Meyer and Wherry (2016) add that the reductions in child mortality due to CHIP are concentrated among the blacks. They find no evidence of CHIP reducing mortality among white children. In contrast, Racine et al. (2001) and Lykens and Jargowsky (2002) find no significant reductions in days of restricted activities among poor children during the early years of Medicaid and CHIP expansions. Similarly, both Dafny and Gruber (2005)

---

eligible households.

[24] Seldon and Hudson (2006) use two sets of instruments: 1). participation in Food Stamp Program and whether parents are covered by employment-based health insurance at the family level, 2). average penetration rates of Medicaid/CHIP at the state level. However, these instruments may be subject to policy endogeneity.

and Currie et al. (2005) find differential results by the different periods of Medicaid/CHIP expansions.

In contrast to previous works that examine the effects of Medicaid and CHIP eligibility, Thompson (2017) finds positive effects of Medicaid and CHIP coverage during the age of 0-18 on adult health outcomes in the long-run. Notably, this study considers adult outcomes, which are largely determined by the earlier years of human capital acquisition.

The literature focusing on educational attainment contains fewer studies. Levine and Schanzenbach (2009) link Medicaid and CHIP eligibility at birth to improvements in reading scores between 4th grade and 8th grade. Likewise, Cohodes et al. (2016) find that childhood Medicaid and CHIP eligibility increases the rates of high school and college completion among poor children. They also find that eligibility at birth as well as through school-age generates educational returns in early adulthood.

In contrast, Cullen et al. (2005) do not find statistically significant effects of CHIP coverage on test scores for elementary children. Extending their study period, we examine whether school-age CHIP coverage over a multi-year period generates educational returns for middle school children.

2.3 Empirical Approach

The primary objective of this paper is to identify the causal effect of the duration of coverage in CHIP on health and education outcomes for middle-school children. Let *i* denote the individual, *j* denote school, and *s* denote state. If CHIP is randomly assigned, then an Ordinary Least Squares (OLS) regression of a health or educational outcome on the duration of CHIP coverage (i.e. the number of years enrolled in CHIP) will yield an unbiased estimate. The OLS regression specification is as follows:

$$HE_{ijs} = \alpha_0 + \alpha_1 \cdot DCHIP_{ijs} + \sigma' \cdot X_{ijs} + \varepsilon_{ijs} \quad (2.1)$$

where HE_{ijs} is the health or educational outcome, $DCHIP_{ijs}$ is the duration of CHIP coverage, X_{ijs} is a vector of control variables containing the individual, school, and state level characteristics, and ε_{ijs} is the error term. In this case α_1 is the effect of CHIP coverage duration on the outcomes.

However, the OLS estimate of α_1 may be biased for several reasons. First, CHIP coverage could be negatively correlated with a child's underlying health. For example, if a sick child is more likely to enroll in CHIP, then the coefficient of α_1 would be downwardly biased. Second, the expansion of CHIP eligibility could be correlated with a state's demographic composition. For instance, if states with healthier populations offer more generous health care plans to children, then children living in these states would achieve better outcomes even in the absence of CHIP. In that case, the OLS estimate would be upwardly biased.

In order to address the endogeneity of CHIP coverage due to individual selection and state demographics, we use simulated eligibility as an instrument, following Currie and Gruber (1996a, 1996b) and Cutler and Gruber (1996). This IV approach accounts for not only selection at the individual level, but also accounts for the fact that a state's demographic composition may be endogenous to CHIP income eligibility. By using a fixed sample to simulate CHIP eligibility, we exploit variation in CHIP rules across states and over time. Critically, we simulate the eligibility on a fixed sample to circumvent the confounding effects from unobservable individual traits and from demographic attributes that could be related to both CHIP coverage and outcomes. Although recent papers corroborate the credibility of using simulated eligibility as the instrument (e.g., Cohodes et al. 2016; Thompson, 2017; Dillender, 2017), we further reduce the potential policy endogeneity by controlling for an extensive set of state-level variables related to children's health and education, and further examine the validity of the simulated eligibility as instrument with a set of falsification tests.

We use the sample of children in their kindergarten year, as a fixed nationally representative cohort that does not vary across states or time. Next, we divide the fixed sample into mutually exclusive demographic cells by the children's household size, race, gender, and age. For each year of the children's school-age, we determine whether an individual child in a specific cell would be eligible for CHIP by comparing household income in the kindergarten year to the eligibility rule of each state and in that specific school-age year. We then calculate the proportion of eligible children for the cell as:

$$Simulated\ Eligibility_c = \frac{1}{n}\sum_{i=1}^{nc} Eligibility_{ci}\ for\ all\ cells \qquad (2.2)$$

where nc is the number of children in the specific cell. Holding the sample in each cell fixed, we calculate the proportion of eligible children in the cell for each specific state-year pair. For example, the simulated eligibility for the fixed cell of female children at age 6 in a white household with 3 family members in Pennsylvania in 2000 is 31% when we apply the CHIP eligibility rule in Pennsylvania in 2000; the simulated eligibility for the same cell in New Jersey in 2002 is 48% when we apply the CHIP eligibility rule in New Jersey in 2002. In this way, we obtain the simulated eligibility measure by the unique household size-race-gender-age cells across all sample states and years. Finally, since our regressor of interest is the duration of CHIP coverage, we construct the duration of simulated eligibility by summing the number of years of simulated eligibility. We use the simulated eligibility duration to instrument for the corresponding CHIP coverage duration of children in a certain demographic cell by their state of residence.

2.4 Data

The Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K), conducted by the National Center for Educational Statistics of the U.S. Department of Education, tracks school experiences of a nationally representative sample of roughly 22,000 kindergarteners (Institute of Education Sciences, 2009).[25] The ECLS-K data collects detailed information on the children, their parents, and school administrators on the child's entry to kindergarten, transition into primary school, and progression through 8th grade. Trained field agents surveyed children in schools and recorded their weight and height. Parents were interviewed on the phone, and school administrators provided information on children's direct and indirect academic performance. Identifiers for state of residence are contained in the restricted-use version of these data. For more information on the ECLS-K, see the User's Manual (Tourangeau et al., 2009).

The information on health insurance varies across the waves. In the spring kindergarten survey (1999), parents were asked whether their child was covered by any insurance.[26] In the following spring waves, parents were asked about the type of the insurance coverage,[27] which includes private health insurance, Medicaid/CHIP, military health insurance, other public health insurance,[28] and no insurance. Our sample includes children who participated in the spring 8th grade wave and had information on CHIP in all school-age spring waves: the spring of first grade (2000); the spring of third grade (2002); the spring of fifth grade (2004); and the spring of eight

[25] The ECLS-K contains respondents in 41 out of 50 states. That excluded states are: Arkansas, Idaho, Montana, Nevada, New Hampshire, North Dakota, South Carolina, Vermont, and West Virginia.

[26] The health insurance question in the kindergarten wave includes private plans that only provide extra cash while an individual is hospitalized, which are not technically health insurance plans.

[27] The original health insurance status does not reconcile coverage among different insurance types. For example, a child's health insurance status can be coded as having CHIP and private plans simultaneously. In some states, CHIP is outsourced to private insurance companies, and carries different names. We classify those reporting both CHIP and private coverage as being covered by CHIP. The share of the sample of those children is around 15%. We then use Hot-Deck Imputation, following the User's Manual (Tourangeau et al., 2009) to fill in the missing values. Our imputed insurance status is in line with that in Cullen et al. (2005).

[28] We exclude the children covered by military insurance plans and other public insurance plans, which reduces the sample by 3%.

grade (2007).[29] Because of the inability to identify CHIP coverage in the kindergarten spring wave, we do not use this wave. Our final sample size is 8,600 children and all models are estimated rely on cross-sectional data of individuals in the 8th grade wave.[30]

Information on health insurance status is collected at a point in time, so the specific time of the transition among different plans is unknown. Therefore, we take the midpoint of the interval between any two waves as the point of transition of insurance plans occurs when children switch insurance plans or enroll in coverage in between survey waves.[31]

We then limit our empirical analysis to the subset of children with household income within 100-300% of FPL in the spring kindergarten (1999) because only children in near-poor to moderate income families were affected by the changes in CHIP income eligibility.[32] This sample contains 2,800 children. In addition, excluding children in families with household income below 100% of FPL mitigates the concern of the confounding effects of other welfare programs.[33]

The ECLS-K contains several measures of health care utilization from the parent surveys. Previous work that assesses the effects of CHIP on access to health care, use whether a child has a usual source of medical care as a measure of access to health care (e.g., Slifkin et al., 2002; Dick et al., 2004). We create an indicator variable of consistent routine care that equals one if the child had routine doctor visits in all waves (i.e., never reported more than year between doctor visits for routine care between 1st grade and 8th grade), and equals zero otherwise. Similarly, we create an

---

[29] In each spring wave, parents were interviewed between March and early July.

[30] Students who do not remain in school are randomly selected to be interviewed in subsequent years at a rate between 44% and 50%. Cross-sectional and longitudinal weights ensure the sample of children in subsequent waves nationally representative.

[31] For example, if a parent reported that the child was covered by private insurance in the 3rd grade, but no insurance in the 5th grade, then we record the duration of private insurance coverage in this interval to be 1 year, or 12 months. In addition, the percentage of children who were enrolled in CHIP but opted out of CHIP in the following wave is below 3%.

[32] During our study period, the lowest CHIP income eligibility is 100% of FPL (e.g., Kansas and Delaware). Children in households with income below the poverty line qualify for Medicaid, and are not affected by the CHIP expansion.

[33] Food Stamp and Temporary Assistance for Needy Families are generally provided to children with household income below 100% of FPL during our study period. More than half of the states in our study period institute the income eligibility of CHIP at 200% of FPL.

indicator variable of consistent dental care use (i.e., never reported more than a year between dental visits between 1st grade and 8th grade). In addition to access to primary health care, we are interested in whether children covered in CHIP were more likely to be diagnosed with illnesses. We create an indicator variable of asthma diagnosis that equals one if the parent had ever been told that their child had asthma up to the 8th grade by physicians.[34] Asthma, one of the most prevalent chronic conditions in children, has disproportionately adverse impacts on lower income children (Akinbami et al., 2009). Since asthma is not preventable but controllable, being diagnosed with asthma in childhood is indicative of better access to care rather than worsening health status. The ECLS-K also provides information on whether the child was diagnosed with sinusitis, hay fever, and diabetes. However, these illnesses are preventable. Being diagnosed with these diseases may be indicative of either better health care utilization or worsening health status. The prior is unknown.

One health outcome measure is parent-evaluated child health status. A parent was asked to rate their child's health on a scale from 1 to 5, indicating the child's perceived health status (1 - excellent, 2 - very good, 3 - good, 4 - fair, 5 – poor). Along with this subjective measure of child health, we have two objective measures: BMI z-score, and an indicator for whether the child was obese (BMI >= 95% of the age-gender adjusted BMI distribution).[35] We measure all health outcomes in the 8th grade since health is a stock variable.

Following the value-added model of Kaestner and Grossman (2009), We use the change in the Item Response Theory (IRT) theta test scores for reading and math between 1st grade to 8th

[34] The consistent routine care binary variable, consistent dental care binary variable, and the indicator of an ever diagnosis of asthma better match to the specification of CHIP coverage duration than the current access to health care, which is a flow variable.
[35] The BMI z-score and weight classification are based on the 2000 release of the CDC Growth Charts (U.S. D.H.H.S, 2002).

grade to measure the child's increase in academic performance.[36] The IRT theta scores were standardized separately to a mean of zero and a standard deviation of one at all waves. A positive difference in the theta scores indicates an improvement in academic performance over time.

Information on the respondent's family, school, and neighborhood is available in the spring 8th grade wave (2007), allowing us to control for individual and school characteristics. The control variables for household socio-demographic characteristics include birth weight, child age (continuous in months), grade, gender, race (white, black, hispanic, and other), population density of residence (urban, suburban, or rural), family income,[37] family size, and the highest education level of the parents (8th grade or less, some high school but did not graduate, high school graduate, some college or 2- year degree, 4-year college graduate, more than 4-year college degree). The control variables for school characteristics include school type (a binary indicator for public school) and the percentage of students at the school eligible for a free or reduced price lunch.

In order to construct the simulated eligibility instrument, we obtain the CHIP income eligibility in all spring waves from a variety of resources.[38] This information is contained in Table 2.1. We construct the duration of simulated eligibility as described in Section 3.

[36] The IRT theta scores are adjusted by the Errata Document provided by IES: http://nces.ed.gov/pubs2012/2012014_Errata.pdf.
IRT scores have two advantages over the raw test scores: 1). IRT scores compensate for the possibility of a low-ability child guessing several difficult items correctly; 2). and make longitudinal measurement of gain in achievement over time possible, though the tests were not identically administered. The IRT scale scores, derived from the IRT theta score, are inappropriate to measure academic gain over time because of its different scales over time. However, the ECLS-K, 1998 does not provide information of how to rescale the IRT scale scores.

[37] Households were asked to report income to the nearest $1,000 income range. We create one measure of family income using the midpoint of the income ranges, a second measure using the upper limit of the income ranges, and a third measure using the lower limit of the income ranges. The empirical results are robust to the three measures. We report empirical results conditional on the midpoint measure for household income.

[38] Our source of CHIP income eligibility are Centers for Medicare and Medicaid Service, Kaiser Family Foundation, Department of Social Services in Virginia, Department of Health in New York State, Department of Social Services in South Dakota, Department of Children & Family Services in Louisiana, and Riley et al. (1998).

While the IV strategy may address the endogeneity problem of individual selection and demographics, this strategy assumes that states do not set CHIP eligibility limits based on average health status and educational outcomes. In order to address the potential issue of policy endogeneity, we control for state-level measures of social economic status and educational resources (SES), including real per capita income, the percentage of adults in the state with a bachelor's degree of higher, the state percentage of overweight and obese girls and boys, the state percentage of overweight adults, the average public school pupil-to-teacher ratio, real total state tax revenue per student, and real state instructional expenditures per teacher. Per capita income and adult education were obtained from the U.S. Census Bureau; the percentage of obese and overweight children we obtained from the Centers for Disease Control and Prevention; and the other state-level characteristics were obtained from the U.S. Department of Education. In addition, we control for the impact of other public welfare programs on educational and health outcomes through state-level measures of participation in the National School Lunch Program (NSLP), School Breakfast Program (SBP), and Summer Food Services Program (SFSP).

Figure 2.1 plots CHIP income eligibility for children above age 6 as of 2007, which is the year of the spring 8th grade wave of ECLS-K. In 2007, thirteen states instituted the CHIP maximum income eligibility above 200% of FPL, including high-income states, such as California, New York, New Jersey, and Massachusetts. Six states had eligibility levels below 200% of FPL. The remaining twenty-two states set their eligibility at 200% of FPL.

Table 2.2 lists the descriptive statistics for insurance status, CHIP eligibility, and outcome variables in all waves. While CHIP eligibility did not dramatically change over years, the share of children enrolled in CHIP steadily rose from 13% in 2000 to 20% in 2007. This suggests a "woodwork effect", where the increase in CHIP coverage is concentrated among those previously eligible for CHIP but not enrolled. Possible explanations include the lag in application simplifications and outreach efforts related to CHIP (Dubay and Kenney, 2009). The share of children covered by private insurance stayed relatively constant over the same period. As a result, the share of uninsured children fell gradually over time.

Table 2.3 lists the descriptive statistics for all control variables at the individual level, school level, and state level.

We apply the child cross-sectional weight in 8th grade to ensure the estimates are nationally representative and cluster the standard errors at the state level.

2.5 Empirical Results

2.5.1 The Effect of Simulated Eligibility on Coverage (1st Stage of IV)

The coefficient of the instrument, simulated eligibility duration from the first stage of our IV model is reported in Table 2.4. The results indicate that simulated eligibility is a statistically significant predictor of an individual's duration of CHIP coverage. Moreover, the F statistic associated with the instrument is 17.38, which is above the conventional threshold of 10 for sufficiently powerful continuous instruments (Stock et al., 2002).

2.5.2 The Effect of CHIP Coverage Duration on Outcomes (2nd Stage of IV)

Table 2.5 presents results for the second stage of the IV-Probit estimation for health care utilization, and results from Probit regressions, for the sake of comparison. The IV coefficients of the effects of CHIP coverage duration on consistent access to routine health care and asthma diagnosis are positive and statistically significant at the 5% level.[39] However, the IV coefficient of consistent dental care is borderline statistically insignificant, though it carries the expected sign.

The IV marginal effects imply that an additional year of CHIP coverage increases the likelihood of receiving consistent routine health care and an asthma diagnosis by 9.14 and 6.96 percentage points, respectively. This is a 16 percent increase in the average probability of receiving consistent routine care and a 51 percent increase in the baseline asthma diagnosis.

The IV effect of CHIP coverage duration on routine care is consistent with Howell and Kenney (2012), who report that one year of Medicaid/CHIP coverage increases the likelihood of receiving a usual source of medical care by 11.8 percentage points. Although the IV coefficient on dental care is insignificant, its magnitude, an annual 6.24 percentage point increase also falls within the interquartile range of the estimated impact of Medicaid/CHIP coverage on dental care (Howell and Kenney, 2012).[40]

The Probit estimates of the impact of CHIP coverage duration on health care utilization are smaller, and are even negative for dental care. Contrasting these with the IV estimates suggests that there is a selection into CHIP by children with poorer baseline access to health care.

We report results for the second stage of the IV estimation for health status and growth in test scores, along with results from OLS in Table 2.6. However, none of the IV coefficients on the CHIP coverage duration are statistically significant.

The imprecisely estimated results for health status and education outcomes are consistent with findings by Cullen et al. (2005). One possible explanation is that children enrolled in CHIP

[39] IV-linear probability model (IV-LPM) generate similar marginal effects as the IV-Probit model, and results of IV-LPM are available upon request.

[40] The interquartile interval of increases in dental care visits due to Medicaid and CHIP coverage ranges from a 4.5 percentage points to a 25 percentage points.

were in good health, leaving little room to improve. Another is that better health and education outcomes may materialize further in the future.

In summary, the IV model finds that longer coverage in CHIP leads to better health care access for middle-school children, though the results for health status and academic performance are imprecisely estimated.

2.5.3 Falsification Tests

Even when we control for a series of observable state characteristics to avoid the potential policy endogeneity, there is possibility that the simulated eligibility duration could be positively correlated with unobserved trends in health at the state level. If children in states with better health care resources are covered by more generous CHIP plans, then the IV estimates would be upward biased. Another concern is that there may be other state-level policies, such as Earned Income Tax Credit (EITC), that are correlated with CHIP and also affect children's health (Case et al., 2002; Currie, 2009). To investigate the validity of the instrument, we conduct two sets of falsification tests.

First, we estimate our IV model on whether the child had access to routine medical care in the kindergarten year. This is not plausibly affected by the duration of CHIP coverage between 1st grade and 8th grade. Results of the probit and IV-probit regression of this binary variable are shown in Table 2.7. The IV coefficient of CHIP school-age duration is not statistically significant, implying that CHIP's expansion is not related to the underlying trend in health care utilization.

Second, since there is no formal way to test the validity of the instrument in a just-identified IV model, we follow the informal approach of Evans and Schwab (1995). Specifically, we include the instrument as a regressor in the Probit model to examine whether the instrument directly impacts the dependent variable, controlling for the independent variables. Again, we focus on the outcomes of health care utilization, given that their IV estimates are statistically significant. The results of the informal test and the original Probit (i.e., without the instrument) are reported in Table 2.8. Neither the coefficient of the CHIP coverage duration varies across the two models, nor the coefficient of the instrument is statistically significant. This suggests no evidence against the exclusion criteria of the instrument.

The results of these falsification tests should be interpreted with caution. A failure to reject the null hypothesis does not prove that the instrument is valid. Rather, the theoretical validity of the instrument is based on research indicating that states institute their own CHIP eligibility to increase public insurance coverage (Currie and Gruber, 1996a, 1996b; Cutler and Gruber, 1996).

2.5.4 Robustness Check

We take the midpoint of the interval between any two waves to measure the duration of CHIP coverage because the specific time of transition among plans is unobservable. We conduct a robustness test for whether our results for the health care utilization are driven by the unobservable timing of transition.

We create one measure of CHIP coverage duration between any two waves using the information on insurance in the leading wave only. For example, if a child was enrolled in CHIP in the 3rd grade, but not in the 5th grade, then we record the duration of CHIP coverage in this interval to be 2 years as opposed to 1 year in our main specification. Likewise, we create another measure of CHIP coverage duration between any two waves using the information on insurance in the trailing wave only.

Results of these two measures of duration, along with the midpoint duration measure (i.e., the main specification) for health care utilization are reported in Table 2.9. All duration specifications generate similar results, suggesting that the unobservable timing of transition among insurance plans does not affect our baseline results.

2.5.5 Subgroup Analysis by Gender

In this section, we investigate whether the effects of duration of CHIP coverage on health care utilization differ by gender. Both Probit and IV-Probit models are estimated separately for boys and girls, and the results are reported in Table 2.10. The first stage F statistics of the IV model for both boys and girls are of similar magnitudes but fall below 10 (F = 8.67 for boys and F = 7.66 for girls).

The results of the IV model in Table 2.10 indicate that the effect of CHIP coverage on health care utilization is statistically significant at the 1% level across all measures for girls, but not for boys. In particular, the girls' IV marginal effect for consistent dental care is statistically significant, implying that an additional year of CHIP coverage increases the likelihood of receiving consistent dental care by 9.10 percentage points, or a 12-percent increase in the baseline probability.[41]

However, we need to interpret the insignificant result of consistent dental care by gender with a caveat because of the low power of the instrument.

2.5.6 Crowd-Out

To investigate the extent of crowd-out, we construct the duration of actual CHIP eligibility by applying CHIP eligibility rules to the actual sample of children in each wave. We IV regress the CHIP coverage duration on the actual eligibility duration, using simulated eligibility duration as the instrument and the same control variables as in the baseline specification. We re-run the IV estimation by substituting the dependent variable with the private insurance coverage duration.

[41] The average rate of receiving consistent dental care for girls in the 8th grade is 0.754.

The results are reported in Table 2.11. The IV coefficients of the actual eligibility duration for both the CHIP coverage duration and the private insurance coverage duration are statistically significant. Taking the ratio of the absolute value of the latter IV coefficient over the former one, we obtain a crowd-out rate of 60%. Our crowd-out estimate reaches the upper bound of the crowd-out rates provided by Gruber and Simon (2008). This is consistent with our higher income range of 100-300% of the FPL then is used some previous studies.[42]

2.5.7 Effects of CHIP Coverage Duration Over Time

Since we study the effects of CHIP coverage duration on health care utilization for middle-school children, it is reasonable to explore the effects of CHIP coverage duration for elementary-school children also. In particular, we focus on the IV estimates of 1st-5th grade CHIP coverage duration on health care utilization among 5th grade children because the instrument, the 1st-5th grade simulated eligibility duration is sufficiently powerful.[43]

We report the effects of CHIP coverage duration for both elementary-school and middle-school children in Table 2.12. As expected, the effects of CHIP coverage duration on consistent routine care and dental care are greater for elementary-school children than for middle-school children. However, the effect of CHIP coverage duration on asthma diagnosis for elementary-school children is statistically insignificant and negative.

[42] Sasso and Buchmueller (2004) focus on the child sample under the 300% of the FPL, including children covered by Medicaid. The child sample in our sample is of higher household income (i.e., 100%-300% of the FPL) which could qualify children for CHIP rather than Medicaid. As a result, our sample is of a higher rate of private insurance and a lower uninsured rate. Gruber and Simon (2008) finds a crowd-out rate of 58% among children in 100%-200% of FPL income households, and a crowd-out rate of 62% among children in 200%-300% of FPL income households. Their study period is 1996-2002.

[43] We apply the child cross-sectional weight in 5th grade in the model for 5th grade children to ensure the estimates are nationally representative. The dependent variables are modified to match the 1st-5th grade duration of CHIP coverage. The 1st-stage F statistic associate with the instrument, the simulated eligibility duration between 1st grade and 5ht grade is 27.12, greater than the conventional threshold of 10. In contrast, the 1st-stage F statistic of the instrument is 0.90 in the model for 3rd grade children, far more below 10.

2.6 Discussion and Conclusion

Recent studies suggest that expansions in the income eligibility of Medicaid and CHIP leads to better health care utilization (Meyer and Wherry, 2016; Bourdreaux et al., 2016), which in turn improves health status (Thompson, 2017), educational attainment (Cohodes et al., 2016), and even pulls up wages in the labor market (Brown et al., 2015). While the estimated effects of the program eligibility, (the "intent to treat effects" (ITT)) are informative, what policymakers would better like to know are the effects of actual coverage, (i.e. the "treatment effects on the treated" (ToT)).[44]

CHIP targets children in near poor and moderate income households that do not qualify for Medicaid. Although these households are not officially defined as poor, many face financial hardships in paying insurance premiums and medical bills. As a result, inadequate access to health care is not uncommon for such households. However, few studies have investigated how CHIP impacts the children in these households. Furthermore, we are not aware of any nationwide study that examines the effects of CHIP coverage in middle school, the stage where investments into child non-cognitive development generate the maximum human capital returns (Cunha et al., 2006). Our study contributes to the current literature on CHIP by providing the first causal effects of the school-age duration of CHIP coverage on a comprehensive set of health and educational outcomes.

We address endogenous individual selection and state demographic composition by using simulated eligibility duration as the instrument. The results of the IV model indicate that children enrolled in CHIP for a longer period of time are more likely to have a consistent source of routine care and their parents are more aware of whether they have asthma.

[44] Since the take-up rate of CHIP is not complete, there is reason to believe that the magnitude of ToT is greater than that of ITT.

Routine care in childhood is believed to be cost effective (e.g., Trunz et al., 2006), reducing the total costs of medical care while improving health. The current framework of CHIP highlights the importance of routine child care. For example, Bright Futures, a national initiative for well children care under the age of 21, advocated by the American Academy of Pediatrics, has been integrated into the Early and Periodic Screening, Diagnosis, and Treatment programs under CHIP programs that are Medicaid extensions.[45]

Although asthma is not curable, early diagnosis of asthma during school-age helps doctors advise a management plan for both parents and schools. This may increase children's attendance because asthma is the leading chronic cause of absenteeism (Akinbami, 2006). However, because few parents provided information on whether their child received asthma treatment conditional on diagnosis, we are unable to link asthma treatment to school attendance.

We find that while CHIP coverage significantly improves utilization of preventive care, we cannot identify an effect of CHIP coverage on health status and academic performance. This is not surprising given the mixed evidence of expanded CHIP eligibility on health and education.[46] Since our measures of health and education are recorded in middle school, we cannot exclude the possibility that CHIP improves health and educational attainment in adulthood, as is reported by Currie et al. (2008) and Cohodes et al. (2016).

One limitation of our analysis is that we lack a powerful instrument for private insurance coverage.[47] Incorporating the duration of private coverage into our framework would shed light

[45] The integration of the Bright Futures program into CHIP is not limited in Medicaid-based CHIP. For instance, the Wyoming department of health also integrates the Bright Futures into its separate CHIP.

[46] For example, Thompson (2017) finds that CHIP eligibility improves adult health, while finds limited evidence of improved access to health care and educational attainment.

[47] We tried the state mandate of employer-sponsored insurance plans and HMO penetration rates to instrument for the private coverage duration. Unfortunately, neither worked.

upon how CHIP affects uninsured children. Another limitation is that we do not have information on CHIP coverage before 2000 when CHIP experienced its initial expansion.[48]

Despite these limitations, the findings in this paper have important policy implications. Although the Affordable Care Act (ACA) includes a maintenance of effort (MOE) provision for CHIP through September 2019, recent proposals have sought to end the federal MOE protection.[49] If the federal MOE is lifted, some states may cut back CHIP when they face budget constraints, shortening children's CHIP coverage duration. For example, in the absence of the federal MOE, Arizona responded to a budget cut by freezing enrollment into its CHIP from January 2010 to May 2012, resulting in a 1 percentage point increase in the uninsured rate among children, in contrast to a nationally declining trend. However, our research shows that CHIP coverage supported by eligibility expansion promotes consistent routine care in childhood that is crucial to healthy development.

Some may argue that cutting back CHIP would not make children worse off because the households could still enroll their children in private plans through the marketplace with tax subsidies. However, the Medicaid and CHIP Payment and Access Commission (MACPAC) reports that this pathway is possible for less than half of the 5.3 million children currently enrolled in separate CHIP. Furthermore, children enrolled in the marketplace would face higher medical costs while receiving less comprehensive benefits than CHIP.

---

[48] However, it is of a less concern given that the rates of CHIP coverage were lower than 10% during its initial two years, 1998 and 1999 (Rosenbach et al., 2007).
https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/rosenbach_2001_5.pdf

[49] The federal MOE protection mandates states to maintain their CHIP and keep the income eligibility as of 2010 in order to make no fewer children eligible for CHIP. The ACA maintenance of efforts (MOE) of CHIP income eligibility is in effect through the FY 2019, although FY 2017 is the last year that federal funding is provided. The MOE institutes different requirements for Medicaid extended CHIP and separate CHIP.

Thus, our results offer some support for extending the current maintenance of CHIP eligibility, but guard against proposed Medicaid block grants as the uncertainty of the ACA grows.

2.7 Tables and Figures

Figure 2.1 State Income Eligibility for CHIP in 2007

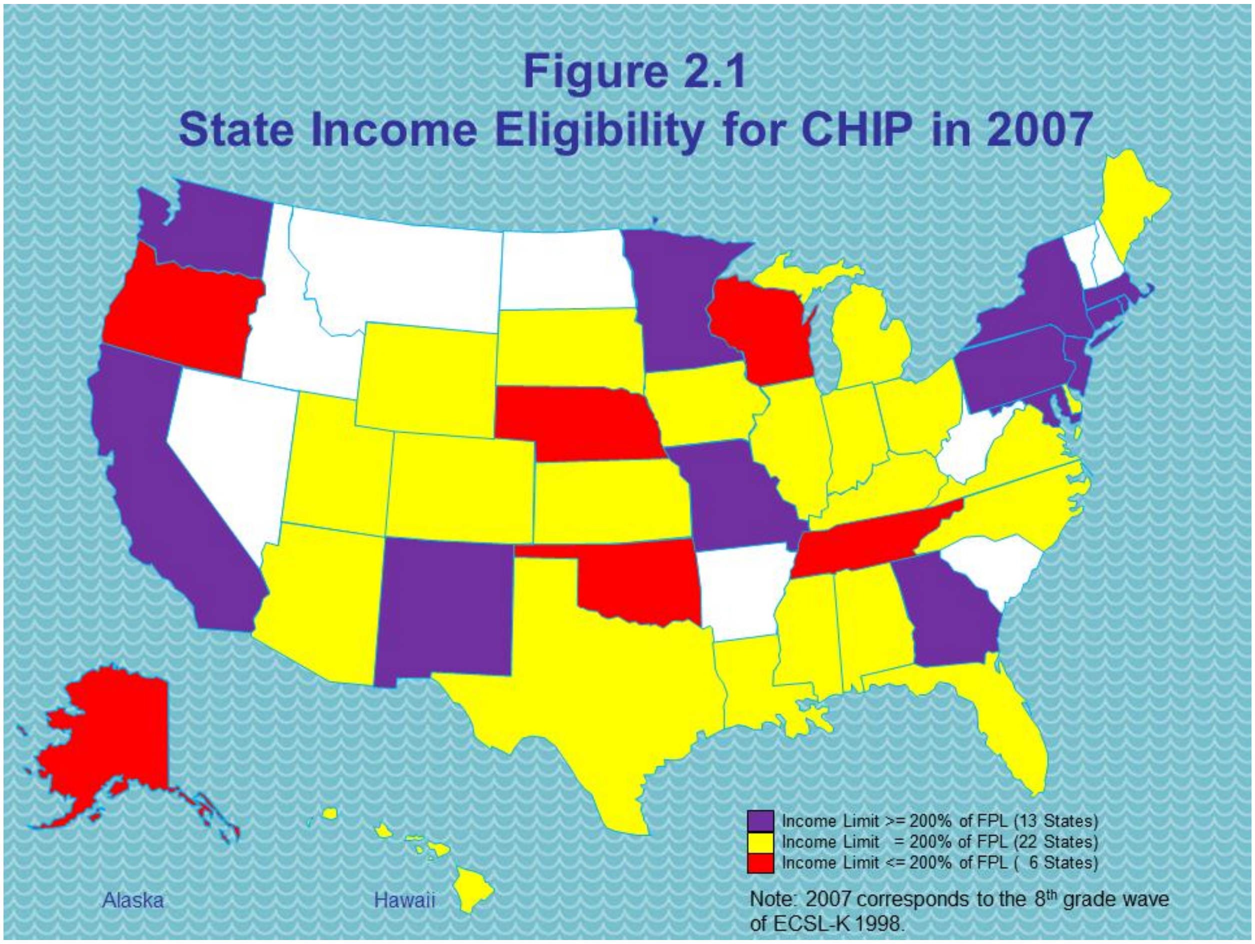

Table 2.1a. Income Eligibility of State CHIP Plans (41 States) Age 1-5

State	Type	Max 1999	Max 2000	Max 2002	Max 2004	Max 2007
Alabama	Medicaid	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Alaska	Medicaid	133%FPL	200%FPL	200%FPL	175%FPL	175%FPL
Arizona	Separate	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
California	Combined	250%FPL	250%FPL	250%FPL	250%FPL	250%FPL
Colorado	Separate	185%FPL	185%FPL	185%FPL	185%FPL	200%FPL
Connecticut	Combined	300%FPL	300%FPL	300%FPL	300%FPL	300%FPL
Delaware	Separate	133%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Florida	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Georgia	Separate	200%FPL	235%FPL	235%FPL	235%FPL	235%FPL
Hawaii	Medicaid	133%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Illinois	Medicaid	150%FPL	150%FPL	185%FPL	200%FPL	200%FPL
Indiana	Medicaid	150%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Iowa	Medicaid	185%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Kansas	Separate	133%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Kentucky	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Louisiana	Medicaid	150%FPL	150%FPL	200%FPL	200%FPL	200%FPL
Maine	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Maryland	Medicaid	200%FPL	200%FPL	300%FPL	300%FPL	300%FPL
Massachusetts	Combined	200%FPL	200%FPL	200%FPL	200%FPL	300%FPL
Michigan	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Minnesota	Medicaid	275%FPL	275%FPL	275%FPL	275%FPL	275%FPL
Mississippi	Medicaid	133%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Missouri	Medicaid	300%FPL	300%FPL	300%FPL	300%FPL	300%FPL
Nebraska	Medicaid	185%FPL	185%FPL	185%FPL	185%FPL	185%FPL
New Jersey	Combined	350%FPL	350%FPL	350%FPL	350%FPL	350%FPL
New Mexico	Medicaid	235%FPL	235%FPL	235%FPL	235%FPL	235%FPL
New York	Separate	192%FPL	222%FPL	250%FPL	250%FPL	250%FPL
North Carolina	Separate	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Ohio	Medicaid	150%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Oklahoma	Medicaid	185%FPL	185%FPL	185%FPL	185%FPL	185%FPL
Oregon	Separate	170%FPL	170%FPL	170%FPL	185%FPL	185%FPL
Pennsylvania	Separate	200%FPL	235%FPL	235%FPL	235%FPL	300%FPL
Rhode Island	Medicaid	250%FPL	250%FPL	250%FPL	250%FPL	250%FPL
South Dakota	Medicaid	140%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Tennessee	Medicaid	No Limit	No Limit	No Limit	133%FPL	133%FPL
Texas	Medicaid	133%FPL	133%FPL	200%FPL	200%FPL	200%FPL
Utah	Separate	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Virginia	Separate	150%FPL	150%FPL	200%FPL	200%FPL	200%FPL
Washington	Separate	200%FPL	250%FPL	250%FPL	250%FPL	250%FPL
Wisconsin	Medicaid	185%FPL	185%FPL	185%FPL	185%FPL	185%FPL
Wyoming	Separate	133%FPL	133%FPL	133%FPL	185%FPL	200%FPL

Table 2.1b. Income Eligibility of State CHIP Plans (41 States) Age 6-16

State	Type	Max 1999	Max 2000	Max 2002	Max 2004	Max 2007
Alabama	Medicaid	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Alaska	Medicaid	100%FPL	200%FPL	200%FPL	175%FPL	175%FPL
Arizona	Separate	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
California	Combined	250%FPL	250%FPL	250%FPL	250%FPL	250%FPL
Colorado	Separate	185%FPL	185%FPL	185%FPL	185%FPL	200%FPL
Connecticut	Combined	300%FPL	300%FPL	300%FPL	300%FPL	300%FPL
Delaware	Separate	100%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Florida	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Georgia	Separate	200%FPL	235%FPL	235%FPL	235%FPL	235%FPL
Hawaii	Medicaid	100%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Illinois	Medicaid	150%FPL	150%FPL	185%FPL	200%FPL	200%FPL
Indiana	Medicaid	150%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Iowa	Medicaid	133%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Kansas	Separate	100%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Kentucky	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Louisiana	Medicaid	150%FPL	150%FPL	200%FPL	200%FPL	200%FPL
Maine	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Maryland	Medicaid	200%FPL	200%FPL	300%FPL	300%FPL	300%FPL
Massachusetts	Combined	200%FPL	200%FPL	200%FPL	200%FPL	300%FPL
Michigan	Combined	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Minnesota	Medicaid	275%FPL	275%FPL	275%FPL	275%FPL	275%FPL
Mississippi	Medicaid	100%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Missouri	Medicaid	300%FPL	300%FPL	300%FPL	300%FPL	300%FPL
Nebraska	Medicaid	185%FPL	185%FPL	185%FPL	185%FPL	185%FPL
New Jersey	Combined	350%FPL	350%FPL	350%FPL	350%FPL	350%FPL
New Mexico	Medicaid	235%FPL	235%FPL	235%FPL	235%FPL	235%FPL
New York	Separate	192%FPL	222%FPL	250%FPL	250%FPL	250%FPL
North Carolina	Separate	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Ohio	Medicaid	150%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Oklahoma	Medicaid	185%FPL	185%FPL	185%FPL	185%FPL	185%FPL
Oregon	Separate	170%FPL	170%FPL	170%FPL	185%FPL	185%FPL
Pennsylvania	Separate	200%FPL	235%FPL	235%FPL	235%FPL	300%FPL
Rhode Island	Medicaid	250%FPL	250%FPL	250%FPL	250%FPL	250%FPL
South Dakota	Medicaid	140%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Tennessee	Medicaid	No Limit	No Limit	No Limit	100%FPL	133%FPL
Texas	Medicaid	100%FPL	100%FPL	200%FPL	200%FPL	200%FPL
Utah	Separate	200%FPL	200%FPL	200%FPL	200%FPL	200%FPL
Virginia	Separate	150%FPL	150%FPL	200%FPL	200%FPL	200%FPL
Washington	Separate	200%FPL	250%FPL	250%FPL	250%FPL	250%FPL
Wisconsin	Medicaid	100%FPL	185%FPL	185%FPL	185%FPL	185%FPL
Wyoming	Separate	100%FPL	133%FPL	133%FPL	185%FPL	200%FPL

Table 2.2. Descriptive Statistics for Health Insurance and Outcome Variables

VARIABLES	Spring of 1st Grade 2000	Spring of 3rd Grade 2002	Spring of 5th Grade 2004	Spring of 8th Grade 2007
Health Insurance and CHIP Eligibility				
Medicaid/CHIP	0.133	0.184	0.217	0.201
	(0.340)	(0.388)	(0.412)	(0.401)
Private Insurance	0.775	0.749	0.728	0.742
	(0.418)	(0.434)	(0.445)	(0.438)
No Insurance	0.092	0.067	0.055	0.057
	(0.289)	(0.250)	(0.228)	(0.231)
Individual Eligibility	0.553	0.54	0.517	0.507
	(0.497)	(0.499)	(0.500)	(0.500)
Simulated Eligibility	0.546	0.573	0.56	0.592
	(0.262)	(0.255)	(0.250)	(0.262)
Duration of CHIP Enrollment in Years	0.133	0.568	0.718	1.408
	(0.340)	(1.140)	(1.321)	(2.411)
Duration of Simulated Eligibility in Years	0.546	1.665	2.253	4.266
	(0.262)	(0.756)	(0.954)	(1.752)
Health and Education				
Consistent Routine Care	0.833	0.702	0.616	0.558
	(0.373)	(0.457)	(0.486)	(0.497)
Consistent Dental Care	0.876	0.818	0.774	0.737
	(0.330)	(0.386)	(0.418)	(0.440)
Asthma Diagnosis	---	0.107	0.123	0.136
	---	(0.309)	(0.329)	(0.343)
Parent-Evaluated Health	1.631	1.647	1.731	1.639
	(0.781)	(0.792)	(0.837)	(0.766)
BMI Z-Score	0.457	0.603	0.679	0.651
	(1.060)	(1.076)	(1.101)	(1.064)
Obesity	0.146	0.198	0.234	0.195
	(0.353)	(0.398)	(0.423)	(0.396)
Reading Theta Score	0.152	0.823	1.069	1.330
	(0.406)	(0.283)	(0.270)	(0.351)
Math Theta Score	0.096	0.749	1.141	1.469
	(0.385)	(0.356)	(0.372)	(0.416)
Observations	2,800	2,800	2,800	2,800

Notes: The cells show the means for the variables in each row. Standard deviations are in parentheses. The survey wave is indicated by the column heading. We subsample to students whose parents provided valid health insurance information and all spring waves and reported household income between 100% through 300% of Federal Poverty Line in the Kindergarten Spring wave. Sample sizes are rounded to the nearest 25 in order to comply with Department of Education non-disclosure requirements for ECLS-K, 1998. The descriptive statistics are based upon unweighted data.

Table 2.3a. Descriptive Statistics for Control Variables at the Individual Level

VARIABLES	Spring of 1st Grade 2000	Spring of 3rd Grade 2002	Spring of 5th Grade 2004	Spring of 8th Grade 2007
Individual Characteristics				
Age in Months	86.976	111.147	134.747	171.448
	(4.416)	(4.415)	(4.435)	(4.412)
Family Incomes ($1,000s)	40.799	45.408	48.947	54.850
	(21.456)	(25.409)	(30.295)	(34.899)
Family Size	4.622	4.632	4.595	4.507
	(1.303)	(1.292)	(1.295)	(1.285)
Female	0.483	0.483	0.483	0.483
	(0.500)	(0.500)	(0.500)	(0.500)
Birth Weight (ounces)	94.547	94.547	94.547	94.547
	(51.867)	(51.867)	(51.867)	(51.867)
White	0.652	0.652	0.652	0.652
	(0.476)	(0.476)	(0.476)	(0.476)
Black	0.180	0.180	0.180	0.180
	(0.384)	(0.384)	(0.384)	(0.384)
Hispanic	0.087	0.087	0.087	0.087
	(0.282)	(0.282)	(0.282)	(0.282)
Other Race	0.081	0.081	0.081	0.081
	(0.273)	(0.273)	(0.273)	(0.273)
Parents' Highest Education	13.823	14.020	14.125	14.230
	(2.195)	(2.243)	(2.258)	(2.290)
Urban	0.338	0.331	0.317	0.279
	(0.473)	(0.471)	(0.466)	(0.449)
Suburban	0.340	0.335	0.321	0.323
	(0.474)	(0.472)	(0.467)	(0.468)
Public School	0.813	0.823	0.831	0.857
	(0.390)	(0.382)	(0.375)	(0.350)
Grade Level of Child	0.976	2.943	4.925	7.917
	(0.159)	(0.242)	(0.274)	(0.300)
% of Students in Free or Reduced-Price Meals in School	35.774	37.913	41.999	41.606
	(23.253)	(26.094)	(21.748)	(21.554)
Observations	2,800	2,800	2,800	2,800

Notes: See the notes to Table 2.2.

Table 2.3b. Descriptive Statistics for Control Variables at the State Level

VARIABLES	Spring of 1st Grade 2000	Spring of 3rd Grade 2002	Spring of 5th Grade 2004	Spring of 8th Grade 2007
State Characteristics				
Real Per Capita Income ($1,000s)	35.392	35.364	36.398	38.230
	(4.253)	(3.966)	(4.067)	(4.829)
% of Overweight Male Children	34.367	34.366	34.366	34.008
	(4.032)	(4.036)	(4.034)	(4.048)
% of Overweight Female Children	26.382	26.377	26.388	28.423
	(4.070)	(4.072)	(4.074)	(4.644)
% of Overweight Adults	36.967	36.940	37.002	36.512
	(1.225)	(0.826)	(1.106)	(1.201)
% of Obese Adults	20.721	22.222	23.164	26.612
	(2.150)	(2.553)	(2.402)	(2.636)
% of Public School Pupil-to-Teacher	16.095	15.863	15.898	15.539
	(2.287)	(2.273)	(2.579)	(2.487)
Real Total Tax Revenues per Student ($1,000s)	9.485	10.085	10.423	11.246
	(1.508)	(1.708)	(1.905)	(2.286)
Real Instructional Expenditures per Teacher ($1,000s)	58.742	60.978	60.926	60.347
	(9.921)	(10.470)	(10.523)	(10.500)
% of Adults with a Bachelor's Degree of Higher	24.217	25.077	26.184	26.520
	(3.739)	(3.766)	(3.817)	(3.916)
% of Students in the National School Lunch Program (NSLP)	59.006	59.841	60.800	63.036
	(11.342)	(11.670)	(11.806)	(11.107)
% of Students in the School Breakfast Program (SBP)	15.384	16.573	17.708	19.916
	(6.567)	(6.789)	(7.181)	(6.897)
% of Students in the Summer Food Services Program (SFSP)	3.914	3.619	3.713	3.648
	(2.369)	(2.250)	(2.715)	(2.771)
Observations	2,800	2,800	2,800	2,800

Notes: See the notes to Table 2.2.

Table 2.4. First Stage: Relationship between Duration of CHIP Coverage and Duration of Simulated Eligibility

VARIABLES	Duration of CHIP Enrollment 1st-8th Grade
Duration of Simulated CHIP Eligibility 1st-8th Grade	0.1750***
	(0.0419)
F Statistic	17.38
Observations	2,700

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Heteroscedasticity-robust standard errors that allow for clustering at the state level are in parentheses. The numbers of observations are rounded to the nearest 25 in order to comply with Department of Education non-disclosure requirements for ECLS-K, 1998. The F statistic corresponds to the hypothesis test that the coefficient for the duration of CHIP enrollment of CHIP eligibility, either simulated or individual, is equal to zero. Additional variables included, but not shown, are: sex, age, grade, race/ethnicity (white, Hispanic, and other race, with black excluded), birth weight, population density (urban or suburban, with rural excluded), the type of school, the percentage of free and reduced-price meals eligible students; family income, family size, parents' highest education; real per capita income in the state, the percentage of adults with a bachelor's degree in the state, the prevalence of childhood obesity among boys and girls in the state, the prevalence of adulthood overweight and obesity, the average pupil/teacher ratio in public schools in the state, real total state tax revenue per student in the state, real state instructional expenditures per student in the state, and the percentages of students in the state participating in the National School Lunch Program, School Breakfast Program, and Summer Food Services Program.
Data: Early Childhood Longitudinal Study, Kindergarten 1998, 8th grade wave (2007).

Table 2.5. Relationship between Duration of CHIP Coverage and Health Care Utilization

VARIABLES	Probit Estimator	IV-Probit Estimator	Probit Marg. Effect	IV-Probit Marg. Effect
Consistent Routine Care	0.0703***	0.2739**	0.0261***	0.0914***
	(0.0168)	(0.1231)	(0.0061)	(0.0332)
Consistent Dental Care	-0.0206	0.1977	-0.0066	0.0628
	(0.0201)	(0.1635)	(0.0064)	(0.0501)
Asthma Diagnosis	0.0637***	0.2984***	0.0143***	0.0696***
	(0.0202)	(0.1073)	(0.0048)	(0.0259)

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: The excluded instrument is the duration of simulated CHIP eligibility 1st grade through 8th grade. We report the estimator and marginal effect of the duration of CHIP enrollment 1st grade through 8th grade. The sample size is 2,700 for the estimation. Estimates from a linear probability model produce similar marginal effects and are available upon request. For additional notes, see Table 2.4.
Data: Early Childhood Longitudinal Study, Kindergarten 1998, 8th grade wave (2007).

Table 2.6. Relationship between Duration of CHIP Coverage and Health Status and Academic Performance

VARIABLES	OLS/Probit Estimator	IV Estimator	OLS/Probit Marg. Effect	IV Marg. Effect
Panel A: Health Status				
Parent-Evaluated Health	0.0383***	0.0303	0.0383***	0.0303
	(0.0122)	(0.1052)	(0.0122)	(0.1052)
BMI Z-Score	-0.0125	-0.1113	-0.0125	-0.1113
	(0.0147)	(0.1032)	(0.0147)	(0.1032)
Obesity	-0.006	-0.0284	-0.0017	-0.0079
	(0.0176)	(0.1804)	(0.0049)	(0.0500)
Panel B: Academic Performance				
Growth in Reading Theta Scores	0.0059	0.0357	0.0059	0.0357
	(0.0064)	(0.0470)	(0.0064)	(0.0470)
Growth in Math Theta Scores	-0.0056	-0.001	-0.0056	-0.001
	(0.0058)	(0.0346)	(0.0058)	(0.0346)

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Parent-Evaluated Health is arranged in a descending scale from 1 to 5. The higher ratings indicate the worse health status. The growth in test scores is measured by the difference of test scores between the 8th grade and 1st grade.

Table 2.7. Falsification Test: Relationship between Duration of CHIP Coverage and Implausibly Affected Outcome

VARIABLES	Probit Estimator	IV-Probit Estimator	Probit Marg. Effect	IV-Probit Marg. Effect
Routine Health Care in Kindergarten	-0.0162	-0.0110	-0.0019	-0.0013
	(0.0290)	(0.2330)	(0.0034)	(0.0274)

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: See notes to Table 2.4.

Table 2.8. Falsification Test: Informal Test of Validity of the Simulated Eligibility

	Consistent Routine Care		Consistent Dental Care		Asthma Diagnosis	
VARIABLES	Probit	Informal Test	Probit	Informal Test	Probit	Informal Test
Duration of CHIP Enrollment 1st-8th Grade	0.0703***	0.0677***	-0.0206	-0.0231	0.0637***	0.0603***
	(0.0168)	(0.0170)	(0.0201)	(0.0201)	(0.0202)	(0.0204)
Duration of Simulated CHIP Eligibility 1st-8th Grade	--	0.0424	--	0.0434	--	0.0516
	--	(0.0298)	--	(0.0438)	--	(0.0328)

Standard errors in parentheses

*** p<0.01, ** p<0.05, *p<0.1

Notes: The informal test is actually an OLS regression of the dependent variable on both the duration of CHIP enrollment and the duration of simulated CHIP eligibility, controlling the same variables as the main model specified in the notes to Table 2.4.

Table 2.9. Robustness Checks: Specifications of Duration of CHIP Coverage

VARIABLES	Probit Estimator	IV-Probit Estimator	Probit Marg. Effect	IV-Probit Marg. Effect
Panel A: Leading Duration Measure				
Consistent Routine Care	0.0626***	0.2685**	0.0233***	0.0901***
	(0.0176)	(0.1244)	(0.0064)	(0.0341)
Consistent Dental Care	-0.0194	0.1883	-0.0062	0.0599
	(0.0206)	(0.1603)	(0.0066)	(0.0494)
Asthma Diagnosis	0.0629***	0.2783**	0.0142***	0.0644**
	(0.0206)	(0.1178)	(0.0049)	(0.0280)
1st Stage F-Statistic		17.05		
Observations		2,700		
Panel B: Trailing Duration Measure				
Consistent Routine Care	0.0721***	0.2718**	0.0267***	0.0900***
	(0.0154)	(0.1213)	(0.0056)	(0.0319)
Consistent Dental Care	-0.0197	0.2042	-0.0063	0.0646
	(0.0189)	(0.1622)	(0.0061)	(0.0492)
Asthma Diagnosis	0.0595***	0.3115***	0.0134***	0.0732***
	(0.0189)	(0.0938)	(0.0045)	(0.0230)
1st Stage F-Statistic		13.96		
Observations		2,700		
Panel C: Midpoint Duration Measure (Main Specification)				
Consistent Routine Care	0.0703***	0.2739**	0.0261***	0.0914***
	(0.0168)	(0.1231)	(0.0061)	(0.0332)
Consistent Dental Care	-0.0206	0.1977	-0.0066	0.0628
	(0.0201)	(0.1635)	(0.0064)	(0.0501)
Asthma Diagnosis	0.0637***	0.2984***	0.0143***	0.0696***
	(0.0202)	(0.1073)	(0.0048)	(0.0259)
1st Stage F-Statistic		17.51		
Observations		2,700		

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: See notes to Table 2.5.

Table 2.10. Subgroup Analysis: Relationship between Duration of CHIP Coverage and Health Care Utilization by Gender

VARIABLES	Probit Estimator	IV-Probit Estimator	Probit Marg. Effect	IV-Probit Marg. Effect
Panel A: Boys				
Consistent Routine Care	0.0672**	0.1726	0.0245**	0.0603
	(0.0292)	(0.2266)	(0.0105)	(0.0728)
Consistent Dental Care	-0.0205	0.0022	-0.0064	0.0007
	(0.0278)	(0.3707)	(0.0087)	(0.1157)
Asthma Diagnosis	0.0731***	0.0753	0.0185***	0.0191
	(0.0245)	(0.2647)	(0.0064)	(0.0660)
1st Stage F-Statistic		8.67		
Observations		1,375		
Panel B: Girls				
Consistent Routine Care	0.0758***	0.3723***	0.0277***	0.1145***
	(0.0223)	(0.0852)	(0.0079)	(0.0180)
Consistent Dental Care	-0.0195	0.2962***	-0.0061	0.0910***
	(0.0271)	(0.0878)	(0.0085)	(0.0250)
Asthma Diagnosis	0.0537**	0.3932***	0.0096**	0.0905***
	(0.0254)	(0.0479)	(0.0047)	(0.0142)
1st Stage F-Statistic		7.66		
Observations		1,300		

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: See notes to Table 2.5.

Table 2.11. Relationship between Duration of Public and Private Insurance Coverage and Duration of Actual Eligibility

VARIABLES	Duration of CHP Enrollment 1st-8th Grade	Duration of Private Enrollment 1st-8th Grade
Duration of Actual CHIP Eligibility 1st-8th Grade	0.3258***	-0.1969**
	(0.0861)	(0.0899)
1st Stage F-Statistic	155.27	
Observations	2,700	

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: The excluded instrument is the duration of simulated CHIP eligibility 1st grade through 8th grade. See the notes to Table 2.4 for additional details of the IV regressions.

Table 2.12. Effects of CHIP Coverage Duration Over Time

VARIABLES	Probit Estimator	IV-Probit Estimator	Probit Marg. Effect	IV-Probit Marg. Effect
Panel A: Duration of CHIP Enrollment 1st-5th Grade				
Consistent Routine Care	0.1102***	0.5212***	0.0384***	0.1684***
	(0.0359)	(0.1619)	(0.0125)	(0.0464)
Consistent Dental Care	-0.0011	0.4534**	-0.0003	0.1350**
	(0.0451)	(0.1783)	(0.0134)	(0.0541)
Asthma Diagnosis	0.0656*	-0.2703	0.0135*	-0.0616
	(0.0347)	(0.2033)	(0.0071)	(0.0542)
1st Stage F-Statistic		27.12		
Observations		2,700		
Panel B: Duration of CHIP Enrollment 1st-8th Grade (Main Specification)				
Consistent Routine Care	0.0703***	0.2739**	0.0261***	0.0914***
	(0.0168)	(0.1231)	(0.0061)	(0.0332)
Consistent Dental Care	-0.0206	0.1977	-0.0066	0.0628
	(0.0201)	(0.1635)	(0.0064)	(0.0501)
Asthma Diagnosis	0.0637***	0.2984***	0.0143***	0.0696***
	(0.0202)	(0.1073)	(0.0048)	(0.0259)
1st Stage F-Statistic		17.51		
Observations		2,700		

Clustering standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Here is the explanation of the dependent variables in Panel A. The indicator variable of consistent routine care equals one if the child had routine doctor visits in spring 1st, 3rd, and 5th grade (i.e., never reported more than year between doctor visits for routine care between 1st grade and 5th grade), or equals zero otherwise. The indicator variable for consistent dental care follows the same rule. The indicator variable of asthma diagnosis that equals one if the child had ever been told to have asthma up to the 5th grade by physicians. See notes to Table 2.5 for additional details.

Chapter 3: The Effect of Participation in School Sports on Academic Achievement Among Middle School Children

3.1 Introduction

Communities in Manhattan, New York City (NYC) scrambled to finance school sponsored sports programs in 2011 after CHAMPS, a pilot program funded by the NYC Department of Education to promote physical activity, had its funding reduced.[50] The cuts forced CHAMPS to supply only one coach per middle school, threatening the sustainability of sports programs at many participating schools. Likewise, in 2014 the state of Pennsylvania announced plans to cut more than $1 billion from the public education system, which led many Pennsylvania middle schools to suspend sports programs due to insufficient budgets.[51] Despite their popularity, schools sports programs are often viewed as expendable when there are shortfalls in educational funding.

The economics literature contains numerous studies on the effect of participation in school sports, either in high school or in college, on educational attainment and labor market outcomes (Long and Caudill, 1991; Maloney and McCormick, 1993; Marsh, 1993; Anderson, 1998; Barron et al., 2000; Robust and Keil, 2000; Eide and Ronan, 2001; Libscomb, 2007; Lozano, 2008; Pfeifer and Cornelissen, 2010; Stevenson, 2010). These papers conclude that students who participate in school sports are less likely to drop out of high school, more likely to attend college, and are more

[50] A DNAINFO New York (10/19/2011) article reports that CHAMPS Middle School Sports and Fitness League cuts funds to middle school sports.
[51] A FactCheck.Org article (06/27/2014) covers the quotes of Tom Wolf.

likely to earn higher wages than their peers who have not participated in sports. However, there are few studies that have investigated the academic return to participating in middle school sports, and none that we are aware of that use nationally representative data.

Human capital accumulation in high school and adulthood builds upon the middle school experience, and there is reason to believe that participation in middle school sports may have a positive impact on intellectual growth and development. Studies indicate that when children enter into the period of adolescence, investments in non-cognitive skills, self-concept,[52] and discipline have a greater impact on long-run human capital accumulation than investments in cognitive skills (Cunha et al., 2006; Heckman et al., 2006; Pfeifer and Reuss, 2008). This is because children with better self-concept and discipline are more efficient at transforming investments in cognitive skills into learning skills. To the extent that socialization and training in sports improve self-concept and discipline, participation in school-sponsored sports may improve learning.

Our paper contributes to the literature on human capital accumulation by providing causal estimates of the effect of participation in school-sponsored sports on academic achievement for middle school children. We analyze data from the nationally representative Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K, 1998). To address the endogeneity of participation in school sports, we apply the method of instrumental variables (IV) to the value-added model of academic performance (Coleman et al., 1966). Results from our IV estimation indicate that the effect of participation in school sports on reading test scores is positive and statistically significant among children in middle school. We find evidence that sports participation is associated with lower levels of absenteeism from class. In addition, we find that

[52] Baumeister (1999) provides a formal definition of self-concept in social psychology: “the individual’s belief about himself or herself, including the personal attributes and who and what the self is.”

the effect of school sports is positive and statistically significant among girls, while not sufficiently identified among boys.

The rest of the paper is organized as follows: we review the benefits of school sports in Section 2. Section 3 contains a description of our empirical approach. We describe the data in Section 4. Section 5 contains our empirical results. In Section 6, we conduct robustness checks of the validity of our instruments and a subgroup analysis by gender. Section 7 concludes the paper.

3.2 Benefits of School Sports

Many papers consistently document positive correlations between school sports and educational attainment. For instance, Marsh (1993) finds a positive correlation between participation in high school sports and college attendance. Long and Caudill (1991) present evidence of a positive correlation between participation in sports at the college level and a higher rate of graduation. These positive correlations are likely mediated through three mechanisms: better academic self-concept, discipline, and health. Figure 3.1 presents a conceptual model, constituting the likely channels through which involvement in school sports affects academic achievement.[53]

3.2.1 Academic Self-Concept and Discipline

A number of research studies establish the way that self-concept, which predicts success in the classroom and the labor market, is improved through school sports. Spady (1970, 1971) proposes that participation in sports increases students' perception of social status, which not only enhances academic ambition and identification within the school, but also facilitates the formation of the

[53] The channels are not exhaustive, but we consider those in Figure 3.1 the most important. There is a possibility that academic self-concept and discipline can be shaped during participation in other after school activities, such as dancing and music classes; however, our data does not contain information on children's participation in other activities in middle school.

skills and attitudes that contribute to a student's future success. Snyder and Spreitzer (1990) propose a role theory in which success in sports provides student athletes with incentives to do better in the classroom.

While a general measure of self-concept is used in most research studies, Marsh et al. (1988) and Marsh (1990) find that academic achievement is significantly correlated with academic self-concept, but not with the general self-concept which encompasses non-academic attributes.

Beyond improvements in academic self-concept, Marks (1977) notes that school sports shape discipline in the classroom within the context of the "spend-and-drain" theory. He argues that participation in sports channels students' abundant energy, making them feel more energetic when doing homework after training for athletic competition. In addition, expanding excess energy in sports stimulates students' interest in school, leading to increased commitment to academic values. As a result, student athletes have lower rates of absenteeism from class than non-athletes.

Kuhn and Weinberger (2005) and Eren and Ozbeklik (2015) add that teamwork and leadership experience in school sports can also contribute to better academic self-concept and discipline.

Both improved academic self-concept and reduced absenteeism, could lead to an improvement in educational outcomes (Marsh, 1993; Spreitzer, 1994). Booth and Gerard (2011) provide the most recent evidence of these pathways by collecting data on students aged 11-12 at four schools in Cleveland. They find positive correlations between academic self-concept and absenteeism in the fall semester and test scores in reading, math, and science in the following spring semester.

3.2.2 Health

Another possible channel through which participation in school sports can influence educational attainment is through its effect on health. Strong et al. (2005) find that more physical activity lowers the probabilities of obesity, cardiovascular disorders, and asthma among school-aged youth. Likewise, Cawley et al. (2013) report that more time spent in physical education (PE) classes reduces the probability of youth obesity, which is a contributor to many health problems in adolescence, such as prediabetes and bone and joint problems (US Department of Health and Human Services, 2010).

However, current research produces mixed evidence on whether the health benefits of physical activity spill over to academic achievement. Strong et al. (2005) identify a moderate association between greater PE class time and gains in academic test scores. In contrast, Cawley et al. (2013) find more time spent in PE class does not affect academic test scores.

3.3 Empirical Approach

If participation in school sports was randomly assigned, one could identify the causal effect of sports participation on academic achievement using the following Ordinary Least Square (OLS) regression:

$$\text{Score} = \alpha_0 + \alpha_1 \cdot Sports + \sigma^T \cdot X + \mu + \varepsilon \tag{3.1}$$

where Score is a student's test score on any one of a number of standardized tests; Sports is a binary indicator for whether the student participated in school-sponsored sports, X is a vector of demographic variables; μ is the student's unobserved endowment of innate ability; and α_1 is the causal effect of interest.

However, the decision to participate in school sports programs is not random. A selection bias may arise due to the fact that participation in school-sponsored sports is optional. For example, if students of lower academic expectations disproportionately participate in school sports, α_1 will have a downward bias.[54] Alternatively, if children of higher socioeconomic status live in school districts that provide higher quality education and greater opportunities to participate in sports, α_1 will have an upward bias.

Only a few studies have attempted to account for this selection. Lipscomb (2006) uses fixed-effect estimation to remove the influence of time-invariant confounding factors that affect both students' propensity to participate in school sports and standardized test scores. However, fixed-effect estimation is not feasible for cross-sectional data. Several studies use the instrumental variable (IV) approach to correct for selection, which is consistent with both time-varying and time-invariant omitted factors. Stevenson (2010) uses the enactment of Title IX in 1972 as an instrument for variation in female athletic participation rates at the state level. She finds that a higher rate of sports participation generates higher rates of college attendance and employment among women. Barron et al. (2000) use the size of schools, library books per capita, and the faculty-to-student ratio as instruments. They find that participation in high school sports results in better educational attainment and higher wages. One limitation of this approach is that the use of school-level information to construct instruments may not account for unobservable attributes of schools that are correlated with higher levels of support for school sports.[55]

[54] Maloney and McCormick (1993) compare the academic achievement of students participating in interscholastic sports with non-participating students by collecting data on all undergraduate students at Clemson University from 1985-1988. They find a positive correlation between low grade point average (GPA) and poor academic background, which is measured by SAT scores and rankings of high schools, among college athletes. In addition, they estimate that approximately 60% of the lower GPA for athletes competing in revenue-generating sports athletes is attributed to the poorer academic background, rather than time management in college.

[55] A similar strategy is exploited by Huang and Humphreys (2012) who use the county-level number of fitness and sports establishment to instrument for the individual probability of participation in physical activity and sports. However, they also control for county-level fixed effects to mitigate the concern over the unobservable county affluence affecting the number of sports establishment.

Edie and Ronan (2001) find a positive effect of high school sports on educational attainment for most students (with the exception of white males) using IV estimation where height at the age of 16 is the instrument. Pfeifer and Cornelissen (2010) also use height as an instrument. They present evidence that participation in high school sports in Germany increases the probability of obtaining a secondary school or professional degree.

There are two requirements for using height as the instrument. First, height is a powerful predictor of school sports participation because children who are relatively tall at a young age have a competitive advantage in many sports, and are more likely to participate as a result (Cordovil et al., 2009).[56]

Second, the instruments must be conceptually valid in that they meet the exclusion restriction. This restriction implies that height cannot have a direct impact on academic achievement. However, previous studies have found a small positive correlation between body height and cognitive ability (Persico et al., 2004; Case and Paxson, 2008).[57]

We are concerned the issue that height could be positively correlated with the unobservable innate ability, invalidating the strategy of using height as the instrument (Rees and Sabia, 2010). Therefore, we consider a standard value-added model:

---

[56] Of course, height is not an advantage in all sports, but the majority of middle school sports are those such as basketball and football for which being taller is good.

[57] Twin studies that attempt to disentangle the influence of environment and genetics on the correlation between height and cognitive skills generate mixed results. Using a large Finnish dataset of 8,798 twin pairs born before 1958 and both alive in 1974, Silventoinen et al. (2000) find that the relationship between height and ability is mostly mediated by environmental factors. When the models are estimated by gender, the small mediating effect of genetics disappears. However, the standard of living in Finland was lower than in western countries until the 1970s. Silventoinen et al. (2004) use a dataset of 5,454 twin pairs living in Minnesota in the 1980s to reexamine the degree to which genetic and environmental factors determine the correlation between body height and education. This study confirms the earlier findings of no genetically-based association between height and education in the United States. In contrast, Sundet et al. (2005) reports that genes explain 35% of the correlation between height and cognition among males using a data of approximately 2,600 Norwegian twins. Likewise, Silventoinen et al. (2006) find that height is positively correlated with cognition when genetics in origin interact with environments using a data of 1,094 Dutch twins. They suggest that height in early childhood (i.e. age 5-7) positively correlates with cognitive skills.

$$\text{Score} = \alpha_0 + \alpha_1 \cdot Sports + \sigma^T \cdot X + \theta \cdot \text{Lagged Score} + \varepsilon \quad (3.2)$$

where the Lagged Score is used to capture the endowment of innate ability. We then derive the differenced specification of value-add model proposed by Kaestner and Grossman (2009) by assuming $\theta = 1$ in equation 2:

$$\text{Score} - \text{Lagged Score} = \text{Growth in Scores} = \alpha_0 + \alpha_1 \cdot Sports + \sigma^T \cdot X + \varepsilon \quad (3.3)$$

In this specification of the value-added model, the unobserved endowment of innate ability, μ is eliminated. In particular, the participation in school sports impacts the growth of test scores, rather than the test score, per se.

We use students' lagged height and the growth in height as instruments, under the assumption that the variation in body height does not directly affect the growth in test scores (i.e., the same influences on test scores in a short period of time). In addition to accounting for selection into school sports, IV estimation has an additional benefit of correcting for classical measurement error (Bound et al., 2002).

3.4 Data

We use the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K, 1998) for our empirical analysis. These data, which are collected by the National Center for Educational Statistics of the U.S. Department of Education, track the school experience of a nationally representative sample of roughly 22,000 children entering kindergarten in the fall of 1998 (Institute of Education Sciences, 2009).[58] The ECLS-K, 1998 contains detailed information on children, parents, and school administrators at entry to kindergarten, transition into primary

[58] The ECLS-K contains respondents in 41 out of 50 states. The excluded states are: Arkansas, Idaho, Montana, Nevada, New Hampshire, North Dakota, South Carolina, Vermont, and West Virginia.

school, and progression through 8th grade. Trained field agents surveyed children in schools and measured their weight and height. Parents were interviewed on the phone, and school administrators provided information on children's direct and indirect academic achievement (Tourangeau et al., 2009).

We limit our sample to middle-school children who were surveyed in the 2007 wave (i.e., the spring of 8th grade for most children) because the question for whether a child participated in school-sponsored sports was only asked in this wave.[59] We create an indicator variable that equals one if the child participated in any school-sponsored sports, either varsity or intramural, in middle school, and equals zero otherwise. Our instrumental variables are the child's height (in inches) in 2000 and the growth in height between 2000 and 2007.[60] We exclude 225 individuals with missing values on school sports participation and height in 2000 and 2007, which results in a final estimation sample of 9,200 students. Sample sizes across different models vary because some dependent variable values are not available for all observations. We apply the child cross-sectional weight in the 2007 wave to ensure that the estimates are nationally representative, and we cluster the standard errors at the Primary Sampling Unit (PSU) level.

The ECLS-K, 1998 contains a direct academic assessment instrument for reading, mathematics, and science. In each subject area, children receive a 10-item routing test in two stages. Performance on certain questions (routing items) in the first stage guided the selection and administration of one of two second-stage (high and low) forms in each subject. The second-stage

---

[59] Nine percent of the students surveyed in the 2007 wave reported being in the 7th grade as opposed to the 8th grade. This is most likely because these children were held back sometime between 2000 and 2007. These students were asked the same question about their participation in school sports and took the same exams in reading, math, and science as students in 8th grade. Estimates of the sub-sample excluding students in the 7th grade are very similar to those using the full-sample. We use the full sample so that the estimates are nationally representative (Tourangeua et al., 2009). Although 3rd and 5th grade waves provided information on participation in sports, no further information was provided for whether these sports were school-based.

[60] We considered other measures of height as instruments, such as the height of the child in 2007, but this instrument set had the highest statistical power.

forms contained items of appropriate difficulty for the level of ability indicated by the routing items. We use change in the Item Response Theory (IRT) theta test scores from 5th grade to 8th grade to measure academic achievement.[61] The IRT theta scores which normally distribute at all waves and range from -3 to 3, are better suited for measuring longitudinal growth than the IRT scale scores (Kieffer, 2011).[62]

In addition, information on the respondent's family, school, neighborhood, and state of residence is available in the 2007 wave, allowing us to use an extensive set of control variables in our empirical model. The most important environmental variable in the control set is birth weight, which captures nutritional intake in utero, which affects both height and cognitive ability (e.g., Black et al., 2007; Xie et al., 2016). Other control variables for household socio-demographic characteristics include child age (continuous in months), grade, gender, race (White, Black, Hispanic, and other), population density of residence (urban, suburban, or rural), family income,[63] family size, and the highest education level of the parents (8th grade or less, some high school but did not graduate, high school graduate, some college or 2- year degree, 4-year college graduate, more than 4-year college degree). The control variables for school characteristics include a binary

[61] The IRT theta scores are adjusted by the Errata Document provided by IES: http://nces.ed.gov/pubs2012/2012014_Errata.pdf.
The IRT scores have two advantages over the raw scores: 1). Using IRT adjusts scores across children who received different second-stage forms, making longitudinal measurement of gain in academic achievement over time possible; 2). IRT scores compensate for the possibility that a low-ability child could guess the answers to several difficult questions correctly.

[62] Kieffer (2011) measures growth of reading ability using ECLS-K, 1998. He suggests that using IRT scale scores may misrepresent growth due to irrelevant factors, such as the number of test items selected for a given ability level at a given grade. In addition, one would need to rescale the IRT scale scores across waves to measure children's academic development. However, the nonlinear rescale algorithm is unknown to researchers. Therefore, the IRT theta scores better represent the developmental growth, corroborated by the ECLS-K user's manual (Tourangeau et al., 2009).

[63] Households were asked to report income to the nearest $1,000 income range. We create one measure of family income using the midpoint of the income ranges, a second measure using the upper limit of the income ranges, and a third measure using the lower limit of the income ranges. The empirical results are robust to the three measures. We report empirical results conditional on the midpoint measure for household income.

indicator for public school and the percentage of students at the school eligible for a free or reduced price lunch.

In the 2007 wave, two teachers (reading and math/science) were asked to rate the student's level of absenteeism from class (i.e., "How often is this student absent from your class?"). The options were: 1 - never, 2 - rarely, 3 - some of the time, 4 - most of the time, 5 - all of the time. We take the average of the response scores for the student's absenteeism from the two teachers as a proxy for the student's academic discipline.[64]

The ECLS-K, 1998 measures each child's self-concept through a Self-Description Questionnaire (SDQ) administered in middle school. Questions in the SDQ are used to measure the extent to which the child internalized his or her problems. Through these questions children were asked how they felt about their academic achievement, including their competence in classroom learning, doing homework, and taking tests. A scale score was calculated as the mean of the scores of the individual items measuring the student's feelings towards his or her academic achievement. This scale score was then rounded to the nearest integer to create the variable SDQ problem internalization. The final measure of SDQ problem internalization variable is ordinal, ranging from 1 to 4 in descending order (1 – strongly agree with the presence of the internalizing problem, 2 – agree, 3 – disagree, 4 – strongly disagree).

The ECLS-K, 1998 also provides measures of children's health status. A parent was asked to rate their child's health on a scale from 1 to 5, indicating health status (1 - excellent, 2 - very good, 3 - good, 4 - fair, 5 – poor).

Table 3.1 contains summary statistics for the ECLS-K variables, including variables used to explore the mechanisms through which participation in school sports impacts academic

[64] Both teachers gave similar ratings on absenteeism in most cases.

achievement, in our empirical models. The participation rate in school sports among middle-school students is 61.1%, highlighting the importance of evaluating the middle school sports programs.

3.5 Empirical Results

3.5.1 Power and Validity of Instruments

The coefficients of the instruments from the first stage of our IV model are reported in Table 3.2. The OLS results indicate that both instruments are statistically significant predictors of an individual's enrollment in school sports. The point estimates of the coefficients on the instruments indicate that being one inch taller in 2000 increases the probability of participating in school sports in middle school by 0.8 percentage points; and growing one inch between 2000 and 2007 translates into an increase in the probability of enrolling in school sports in middle school of 1.3 percentage points. Moreover, the F statistic associated with the instrument set is 11.6, which is above the conventional threshold of 10 for sufficiently powerful continuous instruments (Stock et al., 2002).

Since our IV model is over-identified, we can perform an over-identification test for the validity of the instruments. Hansen's J test examines whether the expected value of the cross product of unobservable errors and functions of observable variables are orthogonal (Hansen, 1982). The null hypothesis is that both instruments, height in 2000, and the growth height between 2000 and 2007, are valid.[65] The chi-square statistics and p-values obtained from the test where the IRT test scores are the dependent variables are shown in Table 3.3. In all cases, we fail to reject the null hypothesis that the instruments are valid.[66]

[65] It is also possible that rejection of the null hypothesis is due to the incorrectly specified conditional moments.

[66] Rees and Sabia (2010) test the validity of height as an IV for sports participation in the Add Health data by regressing residuals from OLS models of academic outcomes (GPA, aspiration to attend college, difficulty in paying attention in class, and difficulty in completing homework) on height and, likewise, find no correlation between height and the residuals.

We rely on the value-added model to difference out the endowment of innate ability. We are concerned that the innate ability could be time varying, invalidating the instruments. We test this possibility by considering fixed-effect regressions of test scores on height, controlling the student's physical activity and the same variables as the baseline model. We use physical activity level as a control variable in this test because it is available in all waves. We create a categorical variable for a student's days of PE class per week reported by teachers (1 - never, 2 – less than once, 3 – 1 or 2 times, 4 – 3 or 4 times, 5 – every day) to measure the student's physical activity. A statistically significant coefficient of height would indicate that height has a direct effect on the growth in test scores. Results of the fixed-effect regressions between any two waves are reported in Table 3.4. The coefficient of height is not statistically significant across any two waves, providing no evidence against the value-added specification.

However, these results should be interpreted with caution. A failure to reject the null hypothesis does not prove that the instruments are valid. Rather, the theoretical validity of the instruments is based on the value-added model where the lagged score captures the endowment of innate ability (Rees and Sabia, 2010). Then, even height could impact test scores differently in early childhood, such a situation is a less concern in our study, however, since the growth in test scores is measured in late childhood (Silventoinen et al., 2006).[67]

3.5.2 The Impact of School Sports on Academic Achievement

Table 3.5 presents estimation results for both the OLS and the IV models of the impact of sports participation on academic achievement. The IV coefficient of the effect of participation in school sports on growth in IRT theta scores in reading from the 5th grade to the 8th grade is positive and

[67] Using Dutch twin data, Silventoinen et al. (2006) report no association between height and cognitive skills in late childhood (i.e., age 10-12), while a positive correlation between height and cognitive skills in early childhood (i.e., age 5-7).

statistically significant at the 5% level. In contrast, IV coefficients for the growth in math and science scores are not statistically significant. The point estimate implies that participation in school sports in middle school improves the reading IRT theta score by 0.27 standard deviations. The OLS estimate of the impact of school sports participation in the growth in reading scores is small and negatively estimated. Contrasting this with the IV estimate suggests that there is selection into score sports by students with lower baseline reading ability. The IV result is smaller than those of Edie and Ronan (2001) who find that participation in high school sports increases the likelihood of attending and graduating from college on the order of 20%-50%, without controlling for children's latent ability.[68] This result differs from Rees and Sabia (2010) who find no impact of high sports participation on GPA in math and English. This discrepancy may be because they do not differentiate school-based sports from other sports programs, or because they use a different measure of academic achievement.[69]

Our results suggest that participation in school sports has a statistically significant impact on the growth of reading scores but no impact on the growth of mathematics and science scores. This is consistent with evidence from White and McTeer (1990) who find that the academic gain from sports participation is more pronounced in subjects where cultural context is more important (such as reading) than in objective subjects (such as math).

3.5.3 Mechanisms

We investigate the plausible mechanisms generating the academic benefits of playing school sports, as described in Figure 3.1. We measure health status using parent evaluated health; we

[68] Edie and Ronan (2001) present empirical results by gender and race. We average their separate IV estimates using the proportions of their sample falling into each racial category in order to derive these estimates for boys and girls.

[69] Rees and Sabia (2010) use children's scores on the Adolescent Health Picture Vocabulary Test to proxy for lagged GPA in math and English in the value-added model.

measure academic discipline using absenteeism from class; and we measure academic self-concept using the SDQ problem internalization instrument that gauges the extent to which children internalize their academic problems.

Table 3.6 displays OLS and IV estimates of the impact of school sports participation on these outcomes. We conduct over-identification tests for each model. The model with parent evaluated health as the outcome fails the test, while the models with absenteeism and SDQ problem internalization pass. As a result, we focus on the OLS estimate for parent's evaluated health, which is qualitatively similar to the IV estimates, but the IV estimates for absenteeism from class and SDQ problem internalization. The OLS estimate, presented in Panel A of Table 3.6, suggests that participation in school sports is positively correlated with better parent-evaluated health status. Consistent with earlier findings that participation in school sports shapes discipline and improves academic self-concept (Marks, 1977; Snyder and Spreitzer, 1990), the IV estimates indicate that participation in school sports leads to lower rates of absenteeism from class and better ratings on SDQ problem internalization.

In order to investigate the next link in the pathway between sports participation and academic achievement, we estimate the impact of parent evaluated health, absenteeism from class, and SDQ problem internalization on the growth in test scores. In this case we present only OLS results because we do not have instruments for these mechanism variables. As a result, the estimates represent associations rather than causal effects because academic outcomes, self-concept, and discipline may be simultaneously determined.[70] Our OLS results in Table 3.7 indicate that reduced

[70] In accordance with symbolic interactionism, Marsh et al. (1993) validate the bilaterally positive association between academic self-concept and academic achievement. On the one hand, higher academic self-concept enhances initiative and facilitates persistence after failure, improving academic achievement. On the other hand, better academic records alleviate depression and worries about school performance.

absenteeism is positively and significantly correlated with the growth in IRT theta scores across all subjects. In contrast, the effects of parent evaluated health status and SDQ problem internalization on test scores are small and not statistically significant. The above results suggest that participation in school sports improves test scores through higher rates of school attendance, but not necessarily through better health and better academic self-concept.

3.6 Robustness Checks

3.6.1 Parental Investment

Positive associations between children's educational attainment and parental human capital investments have been identified in a large number of empirical studies. Keane and Wolpin (2001) find that parental subsidies (e.g., monetary transfers) are the most effective household intervention to encourage young adults at the age of 16 to pursue a postsecondary education. It is possible that children who participate in school sports may receive greater parental investments than those who do not. This could threaten our identification strategy if student athletes benefit from greater parental investments that result in higher test scores. To investigate this possibility, we consider two potential measures of parental investments: whether the child was covered by health insurance, and whether the child had access to routine medical care within the past year. Results of the probit and IV-probit regression of these binary variables on participation in school sports are shown in Table 3.8. Neither of the IV coefficients of participation in school sports is statistically significant, implying that student athletes do not receive more parental investments, as defined by these measures, than do non-athletes.

3.6.2 Discrimination at School

It is also possible that taller children are treated more favorably by teachers and classmates than shorter children in ways that build better self-concept (Persico et al., 2004). The hypothesis of the self-fulfilling prophecy predicts that biased teacher expectations against shorter students ultimately lower their standardized test scores (Jussim, 1991). Likewise, Wentzel (1998) shows that peer support is a positive predictor of students' interest in learning and academic achievement. Therefore, we examine whether favoritism by teachers and classmates at school is related to students' body height. A correlation between height and favoritism would pose a threat to the validity of our instruments because it would imply that height directly improves academic achievement.

Children were asked about whether they felt close to their teachers and classmates (1 – never, 2 – sometimes, 3 – often, 4 - always). We use the degree of closeness to proxy for favoritism at school. We report the results of an OLS regression of these categorical variables on 1st grade height and the growth in height in Table 3.9. None of the OLS coefficients of height in 2000 and the growth in height between 2000 and 2007 is statistically significant in the two OLS models, and the F tests for the joint significance of the variables cannot reject the null hypothesis of no effect of height on the outcomes. These results are not consistent with favoritism towards taller students by their teachers and peers.

3.6.3 Childhood Nutrition and General Self-Concept

One limitation of our data and modeling approach is that we cannot measure the nutritional intake of children throughout their lives, which may impact educational attainment (Strauss and Thomas, 1998). Although disparities in educational outcomes are more likely to be attributed to differences in childhood nutrition in developing countries (e.g., Haddad and Bouis, 1991; Steckel, 1995), we attempt to determine whether differences in nutritional intake impact our results. We do this by estimating a model with an additional control for the nutritional intake of children before middle school. Starting from the 2004 wave (i.e., the spring 5th grade wave) of the ECLS-K, a parent was asked to provide information on the child's vegetable consumption. We create a binary variable for whether the child ate any vegetables other than green salad, potatoes, and carrots on a weekly basis.

Another concern is that taller children may have better self-concept independent of school sports and that this improved self-concept due to height improves test scores (Libscomb, 2007). If this is the case it would violate the IV exclusion restriction. A set of general self-conception (i.e., a collection of beliefs about oneself) questions, adapted from the Rosenberg Self-Esteem Scale (Rosenberg. 1965) are available in the SDQ.[71] These questions asked about the student's perceptions of usefulness, confidence, and pride, independent of academic achievement. Responses to the three questions were standardized separately to a mean of zero and a standard deviation of one. The scale score of the general self-concept is the average of the three standardized scores. A higher score indicates better general self-concept.

Table 3.10 contains the estimates from models that include measures of childhood nutrition and general self-concept. In the IV equation with the additional control for childhood nutrition (Panel A), the estimated coefficients on participation in school sports are very similar to those of our original specification (Panel C) Likewise, including a control for general self-concept in the IV model does not change the coefficient on school sports. This suggests that the academic premium enjoyed by taller students is mediated through participation in school sports. Overall these robustness checks support of the validity of our instruments.

[71] Note that this is a different measure than the one we use for academic self-concept.

3.6.4 Subgroup Analysis: Difference by Gender

In this section, we investigate whether the effect of participation in school sports on academic achievement differs by gender.[72] Both OLS and IV models are estimated separately for boys and girls, and the results are reported in Table 3.11. The first stage F statistic of the IV model for girls is much greater than for boys (F = 29 for girls versus F = 2.7 for boys), suggesting that body height is a powerful predictor of participation in school sports for girls, but not for boys.

The results of the IV model in Table 3.11 indicate that participation in school sports increases the reading scores by 0.15 standard deviations for girls, but there is a statistically insignificant impact for math score and science scores. We also find that participation in school-sponsored sports results in reductions in absenteeism for girls (results available from the authors upon request). However, we need to interpret the result for boys with the caveat that the IV model is not sufficiently identified. Therefore, we cannot rule out the possible benefits of participation in school sports for boys.

3.7 Summary and Conclusions

The prior literature on school sports focuses primarily on the effect of participation in high school sports on educational attainment and labor market outcomes. Past research suggests that academic gains from high-school sports are mediated through increases in the individual's discipline and academic self-concept. However, we are not aware of any nationally representative studies that investigate the academic returns from participation in middle-school sports. To fill this gap in the

[72] The Table of Summary Statistics by Gender will be available upon request. There are only small differences in test scores, participation rates in school sports, and body height and growth on average between boys and girls. For example, the participation rate in school sports is 63% for boys and 59% for girls.

literature, this paper estimates the effect of participation in school sports on academic achievement for children in middle school, using the ECLS-K, 1998. Our analysis is unique among current research on the academic returns to school-sponsored sports for three reasons.

First, we provide the first causal estimates of the academic benefits from playing school sports for children in middle school. We apply IV estimation to value-added specification to correct for endogenous enrollment in school sports due to adverse selection and unobservable environment, using lagged body height and the growth in height as instruments. The results of the IV model indicate that participation in school sports significantly improves the IRT theta scores in reading, but not math and science, for middle-school children.

Second, we explore plausible mechanisms through which sports participation affects academic achievement. We find that the academic return to participation is mediated through better academic discipline.

Finally, we find evidence that the beneficial effect of school sports is significant for girls, but sufficiently identified for boys.

One limitation of our analysis is that we lack information on the length of a child's involvement in school sports in the ECLS-K, 1998. Incorporating the dynamics of participation into the model would shed light upon the optimal duration of involvement in school sports. Another limitation is that we cannot identify different types of programs in middle school, such as intramural sports versus interscholastic sports.

Despite these limitations, the findings in this paper are potentially useful to school administrators. Middle schools in the public education system are subject to federal funding rules instituted under the No Child Left Behind Act (NCLB).[73] Schools that do not meet the NCLB

[73] Eight three percent of middle schools nationwide are public, according to Table 3.1.

requirements for adequate progress in reading and math face fines and sanctions. After the NCLB Act, some schools were compelled to cut funding for physical education and sports to support extra instruction in reading and math (Center on Education Policy, 2006). However, our research shows that cutting funding for sports could be counter-productive.

This paper has other important policy implications. Students generally start participating in school-based sports in middle school because most elementary schools do not offer such programs. According to a Government Accountability Office report (GAO, 2012), opportunities for middle-school students to participate in school sports increased from 2000 through 2006. Nonetheless, budget constraints remain the foremost barrier to middle schools seeking to expand school-based programs to allow broader participation.[74] When local governments face financial difficulties middle schools struggle to maintain school-based sports programs. Some school districts have even instituted "pay-to-play" arrangements, which charge students a participation fee for school sports activities, disadvantaging students from low-income households. In view of the local government's role in promoting children's outcomes in education and closing the gap in long-run human capital accumulation between children of low and high socioeconomic status, policymakers should encourage initiatives that expand access for children to participate in sports within the school context.[75]

[74] The GAO 2012 report cites that purchasing equipment, paying coaches, and subsidizing students' traveling for interscholastic sports are the top 3 financial concerns.

[75] For example, the Mayor's After-School Achievement Program (ASAP) is designed to expand opportunities to sports after school for Houston youth. Through ASAP, the city of Houston funds school agencies to provide after-school programs for middle school youth. ASAP is active during the school year.

3.8 Tables and Figures

Figure 3.1 Conceptual Model of Mechanisms

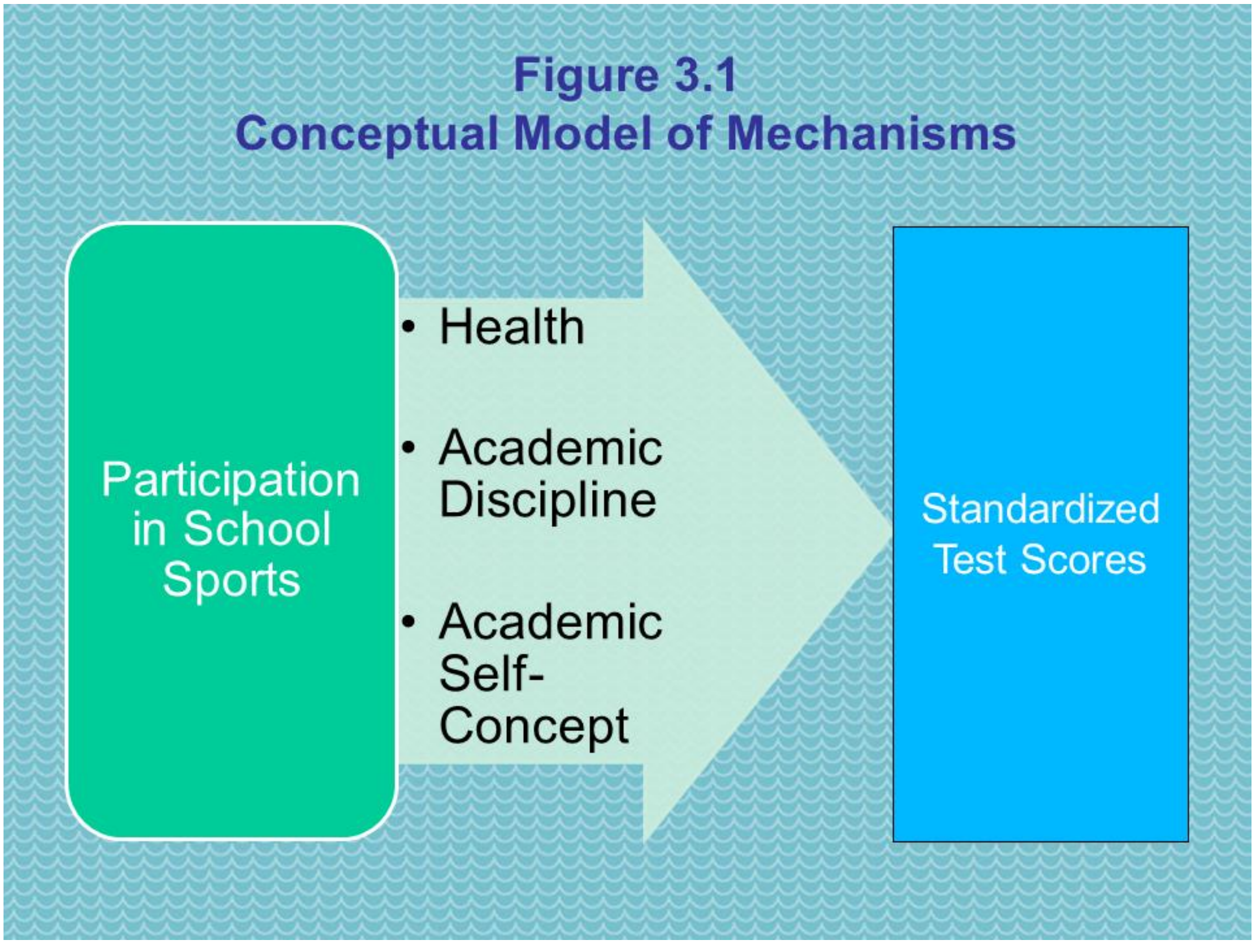

Table 3.1. Descriptive Statistics of Main Estimation in Middle School

VARIABLES	Mean	Std. Dev.
School Sports and Height		
Participation in School Sports	0.610	0.488
Height in 2000 (inches)	48.446	2.397
Growth in Height between 2000 and 2007 (inches)	15.982	2.246
Academic Achievement		
Reading Theta Score	0.288	1.001
Math Theta Score	0.314	1.012
Science Theta Score	0.172	0.964
Student Characteristics		
Age in Months	171.393	4.519
Female	0.497	0.500
8th Grade Child	0.898	0.303
White	0.615	0.487
Hispanic	0.173	0.379
Other Race/Ethnicity	0.108	0.311
Birth Weight (ounces)	91.248	53.545
School Characteristics		
Public School	0.830	0.376
% of Students in Free/Reduced-Price Meals in School	41.479	24.162
Household Characteristics		
Family Incomes ($1,000s)	76.041	59.864
Family Size	4.491	1.298
Parents' Highest Education	13.290	5.084
Urban	0.298	0.457
Suburban	0.361	0.480
Family Rule on Homework	0.938	0.241
Academic *Discipline and Self-Concept, and Health*		
Absenteeism from Class (1-5)	2.093	0.516
SDQ Problem Internalization (1-4)	2.037	0.545
Parent-Evaluated Health (1-5)	1.614	0.774
Observations	9,200	9,200

Notes: Sample sizes are rounded to the nearest 50 and sample means are based upon unweighted data in order to comply with Department of Education non-disclosure requirements for ECLS-K, 1998. Responses to Absenteeism from Class questions are averaged over the reading teacher’s rating and the math/science teacher’s rating (ECLS-K, 1998: User’s Manual, Tourangeua et al., 2009). Scores of Absenteeism from Class are ordered from 1 (never absent) to 5 (absent all of the time). Scores of SDQ Problem Internalization are ordered from 1 (not true at all) to 4 (very true). Scores of Parent-Evaluated Health are ordered from 1 (excellent) to 5 (poor).

Table 3.2. First Stage: Relationship between Height and School Sports Participation

VARIABLES	Participation in School Sports
Growth in Height between 2000 and 2007	0.013***
	(0.004)
Height in 2000	0.008*
	(0.004)
F-Statistic	11.560
Observations	7,500

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: The ECLS-K, 1998 employed a 3-stage probability sample design to select a nationally representative sample of children attending kindergarten in 1998–99. In the first-stage the primary sampling units (PSUs) were geographic areas consisting of counties or groups of counties. The second-stage units were schools within sampled PSUs. The third- and final-stage units were children within schools (ECLS-K, 1998: User's Manual, Tourangeua et al., 2009). Heteroscedasticity-robust standard errors that allow for clustering within PSU are in parentheses. Cross-sectional weights in 8th grade are used to adjust for disproportionate sampling and survey nonresponse. The numbers of observations are rounded to the nearest 50 to comply with non-disclosure requirements for ECLS-K, 1998.
Additional variables included, but not shown, are: sex, age, grade, race/ethnicity (white, Hispanic, and other race, with black excluded), birth weight, population density (urban or suburban, with rural excluded), the type of school, the percentage of free and reduced-price meals eligible students; family income, family size, parents' highest education; real per capita income in the state, the percentage of adults with a bachelor's degree in the state, the average pupil/teacher ratio in public schools in the state, real total state tax revenue per student in the state, and real state instructional expenditures per student in the state.

Table 3.3. Over-Identification Tests: Hansen's J Test

VARIABLES	Reading Theta Score Growth	Math Theta Score Growth	Science Theta Score Growth
Chi2 (1)	1.644	0.011	0.542
P-Value	0.200	0.917	0.462
Observations	7,200	7,250	7,250

*** p<0.01, ** p<0.05, * p<0.1

Notes: We perform Hansen's J Tests for Over-Identification. Hansen's J Tests follow IV-GMM estimation that allows for clustering standard errors on weighted regressions. Degree of freedom of the chi-square test is one. The null hypothesis of Hansen's J tests is that both instruments are valid. Sargan-Basmann Chi-Squared tests following unweighted IV-estimation with homoscedastic standard errors generate similar results. See notes to Table 3.2 for regression details.

Table 3.4. Panel Estimation of the Relationship between Height and Test Theta Scores

	Fixed-Effect Estimation		
VARIABLES	Reading Theta Score	Math Theta Score	Science Theta Score
Panel A: 5th-8th Grade			
Height	-0.002	-0.001	0.006
	(0.002)	(0.002)	(0.004)
Panel B: 3rd-5th Grade			
Height	-0.001	0.003	-0.002
	(0.003)	(0.003)	(0.006)
Panel C: 1st-3rd Grade			
Height	0.006	0.000	-
	(0.004)	(0.004)	-
Panel D: Kindergarten-1st Grade	0.002	-0.001	-
Height	(0.003)	(0.004)	-

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: We add the days of PE class per week into the control set in order to proxy the effect of physical activities on test scores. The fixed-effect estimates are obtained from weighted first difference regressions that allow clustering standard errors. The Hausman test rejects the null hypothesis of random-effect.

Table 3.5. Impact of School Sports Participation on Academic Performance

VARIABLES	Reading Theta Score Growth		Math Theta Score Growth		Science Theta Score Growth	
	OLS	IV	OLS	IV	OLS	IV
Participation in School Sports	-0.021***	0.267**	0.003	-0.059	-0.002	0.116
	(0.006)	(0.132)	(0.006)	(0.117)	(0.016)	(0.218)
1st Stage F-Statistic		16.593		15.813		15.738
Observations	7,200	7,200	7,250	7,250	7,250	7,250

Standard errors in parentheses

*** p<0.01, ** p<0.05, *p<0.1

Notes: Regressions with current height as instrument generate similar results, but with greater standard errors because of the lower F statistic in the 1st stage. Regressions that incorporate the conventional value-added model suggested by equation (3.2) generate similar results. See the notes to Table 3.2 for regression details.

Table 3.6. Investigating the Mechanisms: Relationship between School Sports Participation and Health, Academic Discipline, and Academic Self-Concept

VARIABLES	OLS	IV
Panel A: Health		
Parent-Evaluated Health (lower is better)	-0.148***	-1.151***
	(0.034)	(0.445)
Panel B: Academic *Discipline*		
Absenteeism from Class (lower is better)	-0.055***	-0.335*
	(0.016)	(0.206)
Panel C: Academic Self-Concept		
SDQ Problem Internalization (lower is better)	-0.011	-1.220***
	(0.013)	(0.324)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: In each case, the sample size varies from 6,975 to 7,450. The F-statistics associated with the instruments in the first stage are greater than 10 for all dependent variables, ranging from 11.011 to 11.708. The instruments do not pass the Hansen J test for Parent-Evaluated Health. Parent-Evaluated Health is a categorical variable scaling from 1 to 5 in a decreasing order. Negative marginal effects of Parent-Evaluated Health, Absenteeism from Class, and SDQ Problem Internalization indicate improvement in these measures due to participation in school sports.

Table 3.7. Investigating the Mechanisms: Relationship between IRT Scores and Health, Academic Discipline, and Academic Self-Concept

	OLS Estimation		
VARIABLES	Reading Theta Score Growth	Math Theta Score Growth	Science Theta Score Growth
Parent-Evaluated Health	-0.002	0.002	0.004
	(0.004)	(0.005)	(0.013)
Absenteeism from Class	-0.036***	-0.040***	-0.073***
	(0.005)	(0.004)	(0.012)
SDQ Problem Internalization	0.002	-0.006	0.005
	(0.006)	(0.006)	(0.015)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Parent-Evaluated Health, Absenteeism from Class, and SDQ Problem Internalization are categorical variables where higher values indicate worse outcomes. Negative signs of Parent-Evaluated Health, Absenteeism from Class, and SDQ Problem Internalization indicate positive relationships with test scores.

Table 3.8. Falsification Tests: Relationship between School Sports Participation and Parental Investment

VARIABLES	Health Insurance Coverage		Routine Health Care	
	Probit Marg. Effect	IV-Probit Marg. Effect	Probit Marg. Effect	IV-Probit Marg. Effect
Participation in School Sports	0.002	-0.206	0.053***	0.002
	(0.006)	(0.460)	(0.014)	(0.198)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Both Health Insurance Coverage and Routine Health Care are binary variables. Health Insurance Coverage equals one if the student is covered by any health insurance; Routine Health Care equals one if the student has access to routine health care in middle school.

Table 3.9. Falsification Tests: Relationship between Height and Favoritism at School

VARIABLES	Closeness to Teachers	Closeness to Classmates
Growth in Height between 2000 and 2007	0.003	-0.003
	(0.007)	(0.004)
Height in 2000	-0.004	-0.007
	(0.009)	(0.005)
F-Statistic	0.760	1.470
P-Value	0.473	0.237

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Scores of closeness to teachers and classmates are ordered from 1 (never close) to 4 (always close). Higher values indicate better relationships with teachers and classmates.

Table 3.10. Robustness Checks: Control for Childhood Nutritional Intake and General Self-Concept

VARIABLES	Reading Theta Score Growth		Math Theta Score Growth		Science Theta Score Growth	
	OLS	IV	OLS	IV	OLS	IV
Panel A: Control for Childhood Nutritional Intake						
Participation in School Sports	-0.021***	0.270**	0.003	-0.057	-0.002	0.119
	(0.006)	(0.045)	(0.056)	(0.118)	(0.016)	(0.224)
5th Grade Vegetable Intake	0.013	-0.004	-0.006	-0.004	0.000	-0.011
	(0.013)	(0.015)	(0.012)	(0.017)	(0.020)	(0.027)
F-Statistic		15.057		14.194		14.214
Panel B: Control for General Self-Concept						
Participation in School Sports	-0.024***	0.286**	0.000	-0.068	-0.009	0.007
	(0.007)	(0.143)	(0.006)	(0.126)	(0.016)	(0.223)
Self-Conception	0.025***	-0.004	0.021***	0.025**	0.047***	0.041***
	(0.006)	(0.013)	(0.004)	(0.011)	(0.007)	(0.013)
F-Statistic		16.201		15.653		15.769
Panel C: Primary Specification						
Participation in School Sports	-0.021***	0.267**	0.003	-0.059	-0.002	0.116
	(0.006)	(0.132)	(0.006)	(0.117)	(0.016)	(0.218)
F-Statistic		16.593		15.813		15.738

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: 5th Grade Vegetable Intake is a categorical variable for the times of children ate vegetables per week in 5th grade when survey questions of vegetable consumption first came in availability. It ranges from 1 to 7, with higher values recording more frequent vegetable consumption. The Self-Concept is a continuous variable adapted from the Rosenberg general self-esteem. Its values are standardized to a mean of zero and a standard deviation of one. A higher score indicates better general self-concept.

Table 3.11. Subgroup Analysis: Impacts of School Sports Participation on Academic Achievement by Gender

VARIABLES	Reading Theta Score Growth		Math Theta Score Growth		Science Theta Score Growth	
	OLS	IV	OLS	IV	OLS	IV
Panel A: Male Students						
Participation in School Sports	-0.037	-0.143	0.004	-0.083	0.020	-0.152
	(0.006)	(0.183)	(0.012)	(0.165)	(0.016)	(0.650)
F-Statistic		2.705		2.946		3.187
Observations	3,600	3,600	3,625	3,625	3,700	3,625
Panel B: Female Students						
Participation in School Sports	-0.006	0.145***	-0.001	-0.133	-0.023	0.163
	(0.008)	(0.053)	(0.008)	(0.086)	(0.022)	(0.248)
F-Statistic		29.431		28.307		28.299
Observations	3,625	3,625	3,625	3,625	3,625	3,625

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: The F statistic of IV estimation for boys are far below 10, suggesting imprecise estimates for boys, while the F statistic of IV estimation for girls are above 10. Sample sizes are rounded to the nearest 25.

Bibliography

Adaval, R., and R. S. Wyer Jr. 2011. "Conscious and nonconscious comparisons with price anchors: Effects on willingness to pay for related and unrelated products." *Journal of Marketing Research*, 48(2):355-365.

Akinbami, L. 2006. "The state of childhood asthma, United States, 1980–2005." *Advance Data from Vital and Health Statistics*: No. 381, National Center for Health Statistics.

Akinbami, L., J. Moorman, P. Garbe, and E. Sondik. 2009. "Status of childhood asthma in the United States, 1980-2007." *Pediatrics*, 123(3): 131-145.

Ariely, D., G. Loewenstein, and D. Prelec. 2003. "'Coherent Arbitrariness': Stable demand curves without stable preferences." *Quarterly Journal of Economics*, 118(1):73-106.

Banthin, J., and T. Seldon. 2003. "The ABCs of children's health care: How the Medicaid expansion affected access, burdens, and coverage between 1987 and 1996." *Inquiry*, 40(2): 133-145.

Barron, J., B. Ewing, and G. Waddell. 2000. "The effects of high school athletic participation on education and labor market outcomes." *The Review of Economics and Statistics*, 82(3): 409-421.

Baumeister, R. 1999. *The self in social psychology*, Psychology Press: 195-219.

Black, S., P. Devereux, and K. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *The Quarterly Journal of Economics*, 122 (1): 409-439.

Booth, M., and J. Gerard. 2011. "Self-esteem and academic achievement: a comparative study of adolescent students in England and the United States." *Compare*, 41(5): 629-648.

Boudreaux, M., E. Golberstein, and D. McAlpine. 2016. "The long-term impacts of Medicaid exposure in early childhood: evidence from the program's origin." *Journal of Health Economics*, 45 (1): 161–175.

Bound, J., C. Brown, and N. Mathiowetz. 2002. "Measurement error in survey data." In: Heckman, J., Leamer, *Handbook of Econometrics*, Springer, Vol. 5: 3705-3843.

Brown, D., A. Kowalski, and I. Lurie. 2015. "Medicaid as an investment in children: What is the long-term impact on tax receipts?" *NBER Working Paper*, No.20835.

Case, A., D. Lubotsky, and C. Paxson. 2002. Economic status and health in childhood: The origins of the gradient." *American Economic Review*, 92(5): 1308-1334.

Case, A., and C. Paxson. 2008. “Stature and status: height, ability, and labor market outcomes.” *Journal of Political Economy*, 116(3): 499–532.

Cawley, J., D. Frisvold, and C. Meyerhoefer. 2013. “The impact of physical education on obesity among elementary school children” *Journal of Health Economics,* 32(4): 743-755.

Clemens, J. 2015. “Regulatory redistribution in the market for health insurance.” *American Economic Journal: Economic Policy*, 7(2): 109–134.

Cohodes, S., D. Grossman, S. Kleiner, and M. Lovenheim. 2016. “The effect of child health insurance access on schooling: Evidence from public insurance expansions.” *Journal of Human Resources*, 51(3): 727-759.

Coleman, J., E. Campbell, C. Hobson, F. McPartland, A. Mood, and F. Weinfeld. 1966. “Equality of educational opportunity.” National Center for Education Statistics, Washington, DC.

Cordovil, R., D. Araujo, K. Davids, L. Gouveia, J. Barreiros, O. Fernandes, and S. Serpa. 2009. “The influence of instructions and body-scaling as constraints on decision-making processes in team sports.” *European Journal of Sports Science*, 9(3): 169-179.

Cullen, J., P. DeCicca, and C., Volden. 2005. “The impact of state CHIP programs on early childhood health insurance coverage, utilization and outcomes.” *Economic Research Initiative on the Uninsured*, University of Michigan, Ann Arbor.

Cunha, F., J. Heckman, and L. Lochner. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In: Hanushek, E., and F. Welch, *Handbook of the Economics of Education*, Elsevier, Vol. 1: 709-747.

Currie, J., and J. Gruber. 1996a. “Health insurance eligibility, utilization of medical care, and child health.” *Quarterly Journal of Economics*, 111(2): 431-466.

Currie, J., and J. Gruber. 1996b. “Saving babies: the efficacy and cost of recent changes in the Medicaid eligibility of pregnant women.” *Journal of Political Economics,* 104(6): 1263-1296.

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. 2009. “Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development.” *Journal of Economic Literature,* 47 (1): 87–122.

Cutler, D., and J. Gruber. 1996. “Does public insurance crowd out private insurance.” *Quarterly Journal of Economics*, 111(2): 391-430.

Dafny, L., and J. Gruber. 2005. "Public insurance and child hospitalizations: access and efficiency effects." *Journal of Public Economics,* 89(1): 109-129.

Davidoff, A., G. Kenney, and L. Dubay. 2005. "Effects of the state Children's Health Insurance Program expansions on children with chronic health conditions." *Pediatrics*, 116(1): 34-42.

Dick, A., C. Brach, R. Allison, E. Shenkman, L. Shone, P. Szilagyi, J. Klein, and E. Lewit. 2004. "SCHIP's impact in three states: How do the most vulnerable children fare?" *Health Affairs*, 23(5): 63-75.

Dillender, M. 2017. "Medicaid, family spending, and the financial implications of crowd-out." *Journal of Health Economics*, 53: 1–16.

Dubay, L., and G. Kenney. 2009. "The impact of CHIP on children's insurance coverage: an analysis using the National Survey of America's Families." *Health Service Research*, 44(6):2040-59.

Duan, N. 1983. "Smearing estimate: A nonparametric retransformation method." *Journal of the American Statistical Association*, 78(383):605-610.

Duan, N., W. G. Manning Jr, C. Morris, and J. Newhouse. 1983. "A comparison of alternative models for the demand for medical care." *Journal of Business & Economic Statistics*, 1(2): 115-126.

Edie, E., and N. Ronan. 2001. "Is participation in high school athletics an investment or a consumption good? Evidence from high school and beyond." *Economics of Education Review*, 20(5): 431-442.

Eren, O., and S. Ozbeklik. 2015. "Leadership Skills and Wages Revisited: Is There a Causal Relation?" *Journal of Human Capital,* 9(1): 45-63.

Erikson, E. 1968. *Identity: Youth and Crisis,* W. W. Norton & Company, 93-95.

Evans, W., and R. Schwab. 1995. "Finishing high school and starting college: do catholic schools make a difference." *Quarterly Journal of Economics*, 110(4): 941–974.

Fox, M., J. Moore, R. Davis, and R. Heintzelman. 2003. "Changes in reported health status and unmet need for children enrolling in the Kansas Children's Health Insurance Program." *American Journal of Public Health*, 93(4): 579-582.

Frean, M., J. Gruber, and B. Sommers. 2017. "Premium subsidies, the mandate, and Medicaid expansion: Coverage effects of the Affordable Care Act." *Journal of Health Economics*, 53: 72-86.

Gergaud, O., K. Storchmann, and V. Verardt. 2015. "Expert opinion and product quality: Evidence from New York City restaurants." *Economic Inquiry*, 53(2):812-835.

Grossman, M. 1972. "On the concept of health capital and the demand for health." *Journal of Political Economy*, 80(2): 223-255.

Gruber, J., and K. Simon. 2008. "Crowd-out 10 years later: Have recent public insurance expansions crowded out private health insurance?" *Journal of Health Economics*, 27(2): 201-217.

Hackmann, M., J. Kolstad, and A. Kowalski. 2015. "Adverse selection and an individual mandate: when theory meets practice." *American Economic Review*, 105(3):1030–1066.

Haddad, L., and H. Bouis. 1991. "The impact of nutritional status on agricultural productivity: wage evidence from the Philippines." *Oxford Bulletin of Economics and Statistics*, 53(1): 45–68.

Ham, J., 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.

Hansen, L. 1982. "Large sample properties of generalized method of moments estimators." *Econometrica*, 50(4): 1029-1054.

Heckman, J. 1979. "Sample selection bias as a specification error." Econometrica, 47(1):153-161.

Heckman, J., J. Stixrud, and S. Urzua. 2006. "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor Economics*, 24(3): 411-482.

Howell, E., S. Decker, S. Hogan, A. Yemanem, and J. Foster. 2010. "Declining child mortality and continuing racial disparities in the era of the Medicaid/SCHIP insurance coverage expansions." *American Journal of Public Health*, 100(12): 2500-2506.

Howell, E., and G. Kenney. 2012. "The impact of the Medicaid/CHIP expansions on children: A synthesis of the evidence." *Medical Care Research and Review*, 69 (4): 372–396.

Huang, H., and B. Humphreys. 2012. "Sports participation and happiness: Evidence from US microdata." *Journal of Economic Psychology*, 33(4): 776-793.

Jaeger, D. A., and K. Storchmann. 2011. "Wine retail price dispersion in the united states: Searching for expensive wines." *American Economic Review*, 101(3):136-141.

Jussim, L. 1991. "Social perception and social reality: A reflection-construction model." *Psychological Review*, 98(1): 54–73.

Kaestner, R., and M. Grossman. 2009. "Effects of weight on children's educational achievement." *Economics of Education Review*, 28(6): 651–661.

Keane, M., and K. Wolpin. 2001. “The effect of parental transfers and borrowing constraints on educational attainment.” *International Economic Review*, 42(4): 1051-1103.

Kieffer, M. 2011. “Converging trajectories: reading growth in language minority learners and their classmates, kindergarten to grade eight.” *American Educational Research Journal*, 48(5): 1187–1225.

Kuhn, P., and C. Weinberger. 2005. “Leadership Skills and Wages.” *Journal of Labor Economics*, 23(3): 395-436.

Lee, H., W. Tian, and A. Tomohara. 2008. “The State Children's Health Insurance Program: participation and substitution.” *The Social Science Journal*, 45(3): 382-400.

Levine, P., and D. Schanzenbach. 2009. “The impact of children's public health insurance expansions on educational outcomes.” Forum for Health Economics & Policy, Berkeley Electronic Press, vol. 12(1).

Libscomb, S. 2007. “Secondary school extracurricular involvement and academic achievement: a fixed effects approach.” *Economics of Education Review*, 26(4): 463-472.

Long, J., and S. Caudill. 1991. “The impact of participation in intercollegiate athletics on income and graduation.” *Review of Economics and Statistics*, 73(3): 525-531.

Lozano, F. 2008. “Language, high school leadership and the postsecondary outcomes of Hispanic students.” *Economics of Education Review*, 27(3): 342-353.

Lykens, K., and P. Jargowsky. 2002. “Medicaid matters: Children’s health and Medicaid eligibility expansions.” *Journal of Policy Analysis and Management*, 21(2): 219-238.

Maloney, M., and R. McCormick. 1993. “An examination of the role that intercollegiate athletic participation plays in academic achievement.” *The Journal of Human Resources*, 28(3): 555-570.

Manning, W. G., N. Duan, and W. Rogers. 1987. “Monte carlo evidence on the choice between sample selection and two-part models.” *Journal of Econometrics*, 35(1):59-82.

Marks, S. 1977. “Multiple Roles and Role Strain: Some Notes on Human Energy, Time and Commitment.” *American Sociological Review*, 42(6): 921-936.

Marsh, H., B. Byme, and R. Shavelson. 1988. “A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement.” *Journal of Educational Psychology*, 80(3): 366-380.

Marsh, H. 1990. “A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification.” *Educational Psychology Review*, 2(2): 77-172.

Marsh, H. 1993. “The effects of participation in sport during the last two years of high school.” *Sociology of Sport Journal*, 10(1): 18-43.

Meyer, B., and L. Wherry. 2016. “Saving teens: Using a policy discontinuity to estimate the effects of Medicaid eligibility.” *Journal of Human Resources* (press).

Persico, N., A. Postlewaite, and D. Silverman. 2004. “The effect of adolescent experience on labor market outcomes: the case of height.” *Journal of Political Economy*, 112(5): 1019-1053.

Pfeifer, C., and K. Reuss. 2008. “Age-dependent skill formation and returns to education.” *Labour Economics*, 15(4): 631-646.

Pfeifer, C., and T. Cornelissen. 2010. “The impact of participation in sports on educational attainment - New evidence from Germany.” *Economics of Education Review*, 29(1): 94-103.

Racine, A., R. Kaestner, T. Joyce, and G. Colman. 2001. “Differential impact of recent Medicaid expansions by race and ethnicity.” *Pediatrics*, 108(5): 1135-1142.

Rees, D., and J. Sabia. 2010. “Sports participation and academic performance: evidence from the National Longitudinal Study of Adolescent Health.” *Economics of Education Review*, 29(5): 751-759.

Riley, T., C. Pernice, and R. Mollica. 1998. “How will states implement Children’s Health Insurance Programs?” *Health Affairs*, 17(3): 260-263.

Robust, J., and J. Keil. 2000. “The relationship between athletic participation and academic performance: evidence from NCAA Division III.” *Applied Economics*, 32(5): 547-558.

Rosenbach, M., C. Irvin, A. Merrill, S. Shulman, J. Czajka, C. Trenholm, S. Williams, S. Limpa-Amara, and A. Katz. 2007. “National evaluation of the State Children’s Health Insurance Program: A decade of expanding coverage and improving access.” Centers for Medicare and Medicaid Services. Baltimore, MD.

Rosenberg, M. 1965. *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press, 16-23.

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

Seldon, T., J. Hudson, and J. Banthin. 2004. “Tracking changes in eligibility and coverage among children.” *Health Affairs*, 23(5): 39-50.

Seldon, T., and J. Hudson. 2006. “Access to care and utilization among children: Estimating the effects of public and private coverage.” *Medical Care*, 44(5): I19-I26.

Silventoinen, K., J. Kaprio, and E. Lahelma. 2000. “Genetic and environmental contributions to the association between body height and educational attainment: a study of adult Finnish twins.” *Behavior Genetics*, 30(6):447-485.

Silventoinen, K., R. Krueger, T. Bouchard, J. Kaprio, and M. McGue. 2004. “Heritability of body height and educational attainment in an international context: comparison of adult twins in Minnesota and Finland.” *American Journal of Human Biology*, 16(5):544-555.

Silventoinen, K., D. Posthuma, T. Van Beijsterveldt, M. Bartels, and D. Boomsma. 2006. “Genetic contributions to the association between height and intelligence: evidence from Dutch twin data from childhood to middle age.” *Genes, Brain and Behavior* 5(8): 585–595.

Slifkin, R., V. Freeman, and P. Silberman. 2002. “Effect of the North Carolina State Children's Health Insurance Program on beneficiary access to care.” *Archives of Pediatric & Adolescent Medicine*, 156(12): 1223-1229.

Snyder, E., and E. Spreitzer. 1990. "High school athletic participation as related to college attendance among Black, Hispanic, and White males: A research note.” *Youth and Society*, 21(3): 390-398.

Spady, E. 1970. “Lament for the letterman: effects of peer status and extracurricular activities on goals and achievement.” *American Journal of Sociology*, 75(4): 680-702.

Spady, E. 1971. “Status, achievement, and motivation in the American high school.” *School Review*, 79(3) 379-403.

Spreitzer, E. 1994. “Does participation in interscholastic athletics affect adult development? A longitudinal analysis of an 18–24 age cohort.” *Youth and Society*, 25(3): 368–387.

Steckel, R. 1995. “Stature and the standard of living.” *Journal of Economic Literature*, 33(4): 1903–1940.

Stevenson, B. 2010. “Beyond the classroom: Using Title IX to measure the return to high school sports.” *The Review of Economics and Statistics*, 92(2): 284-301.

Stock, H., J. Wright, and M. Yogo. 2002. “A survey of weak instruments and weak identification in generalized method of moments.” *Journal of Business and Economic Statistics*, 20(4): 518-529.

Strauss, J., and D. Thomas. 1998. “Health, nutrition, and economic development.” *Journal of Economic Literature*, 36(2): 766-817.

Strong, W., R. Malna, C. Blimkie, S. Daniels, R. Dishman, B. Gutin, A. Hergenroeder, A. Must, P. Nixon, J. Pivarnik, T. Rowland, S. Trost, and F. Trudeau. 2005. “Evidence based physical activity for school-age youth.” *Journal of Pediatrics*, 146(6): 732-737.

Sundet, J., K. Tambs, J. Harris, P. Magnus, and T. Torjussen. 2005. "Resolving the genetic and environmental sources of the correlation between height and intelligence: a study of nearly 2600 Norwegian male twin pairs." *Twin Research and Human Genetics*, 8(4): 307–311.

Thompson, O. 2017. "The long-term health impacts of Medicaid and CHIP." *Journal of Health Economics*, 51: 26-40.

Tourangeau, K., C. Nord, T. Lê, A. Sorongon, and M. Najarian. 2009. "Early Childhood Longitudinal Study, Kindergarten Class of 1998 – 99 (ECLS-K), combined user's manual for the ECLS-K eighth-grade and K-8 Full sample data files and electronic codebooks (NCES 2009-004)." National Center for Education Statistics, Washington, DC.

Trunz, B., P. Fine, and D. Dye. 2006. "Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness." *The Lancet*, 367(9517): 1173–1180.

Wang, H., E. Norton, and R. Rozier. 2007. "Effects of the State Children's Health Insurance Program on access to dental care and use of dental services." *Health Services Research*, 42 (4): 1544-1563.

White, P., and W. McTeer. 1990. "Sport as a component of cultural capital: Survey findings on the impact of participation in different sports on educational attainment in Ontario high schools." *Physical Education Review*, 13(1):66-71.

Wentzel, K. 1998. "Social relationships and motivation in middle school: The role of parents, teachers, and peers." *Journal of Educational Psychology*, 98(2): 202-209.

Xie, Z., S. Chou, and J. Liu. 2016. "The short-run and long-run effects of birth weight: Evidence from large samples of siblings and twins in Taiwan." *Health Economics*, Forthcoming. DOI: 10.1002/hec.3367.

Zhang, J., and M. Wedel. 2009. "The effectiveness of customized promotions in online and offline stores." *Journal of Marketing Research*, 46(2):190-206.

2006 *Ten big effects of the No Child Left Behind Act on public schools.* Center on Education Policy, The George Washington University, Washington, D.C.

2010 *The surgeon general's vision for a healthy and fit nation.* U.S. Department of Health and Human Services, Office of the Surgeon General, Rockville, MD.

2012 *K-12 Education: school-based physical education and sports program.* United States Government Accountability Office, Washington, D.C.

Curriculum Vitae

(May 2, 2017)

Xiaohui (Ronnie) Guo

Rauch Business Center
621 Taylor Street
Bethlehem, PA 18015
xig310@lehigh.edu
`http://www.lehigh.edu/~xig310/`

Education

Lehigh University — 2012-2017
Ph.D. in Economics

Lehigh University — 2010-2012
M.S. in Economics

University of International Business and Economics — 2006-2010
B.A. in Finance

Research Fields

Primary Field: Health Economics
Secondary Fields: Applied Econometrics, Industrial Organization

Publications

1. "The Long-run and Short-run Impacts of Urbanization on Carbon Dioxide Emissions", *Economic Modelling*, 53, 2016, pp. 208-215 (SSCI), (with P. Sheng).

2. "Coal, Oil, and Clean Energy: Which Contributes Most to the Low Energy Efficiency in China", *Utilities Policy*, 35, 2015, pp. 67-71 (SSCI), (with P. Sheng, W. Xie).

Working Papers

1. Guo, X., C. Meyerhoefer, "The Effect of Participation in School Sports on Academic Achievement Among Middle School Children".

2. J. Dearden, Guo, X., C. Meyerhoefer, "Restaurant Wine Pricing".

3. Guo, X., C. Meyerhoefer, "The Effect of Coverage through State Children's Health Insurance Program on Health and Academic Performance of Middle School Children".

4. P. Sheng, Y. He, Guo, X., "The Impact of Urbanization on Energy Consumption and Efficiency".

Work in Progress

1. S. Sherer, C. Meyerhoefer, S-Y. Chou, M. Deily, Guo, X., J. Chen "Provider Satisfaction with Electronic Health Records and the Implications for Patient Perceptions of Care Quality".

Research Grants

1. National Statistical Science Fund, China, Grant No. 2015073 (2015-present)

Investigation of the New Normal in Coordination Between Economic Development and Environment

PI: P. Sheng Role: Co-Investigator

2. Doctoral Travel Grants for Global Opportunities, Lehigh University (September-December 2016)

Instructor Experience

Intermediate Microeconomics: Summer, Fall 2016; Lehigh University

Statistical Methods: Spring 2017; Lehigh University

Conferences and Seminars

Conference Presentations

J. Dearden, Guo, X., C. Meyerhoefer, "Restaurant Wine Pricing"

- *Eastern Economics Association 41th Annual Conference, New York City, NY, February 2015*
- *Chinese Economists Society North American Annual Conference, University of Michigan, Ann Arbor, MI, March 2015*
- *14th Annual International Industrial Organization Conference, Drexel University, Philadelphia, Pennsylvania April 2016*

Guo, X., C. Meyerhoefer, "The Effect of Coverage through State Children's Health Insurance Program on Health and Academic Performance of Middle School Children"

- *Eastern Economics Association 42th Annual Conference, Washington D.C., February 2016*
- *Chinese Economists Society North American Annual Conference, Sacramento, CA, April 2016*
- *6th Biennial Conference of the American Society of Health Economists, University of Pennsylvania, Philadelphia, PA, April 2016*

Guo, X., C. Meyerhoefer, "The Effect of Participation in School Sports on Academic Achievement Among Middle School Children"

- *Southern Economics Association 86th Annual Meetings, Washington D.C., November 2016*
- *8th International Symposium on Human Capital and Labor Markets, Central University of Finance and Economics, Beijing, China, December 2016*

Professional Certificates

1. Chartered Financial Analyst (CFA) − Passed Level 3, CFA Institute.
2. Financial Risk Manager (FRM) − Passed Level II, GARP.
3. SAS Certified Advanced Programmer, SAS Global Certification Program.

Professional Service

Referee for *Agricultural and Resource Economics Review*.

Awards

1. Teaching Assistant of the Year, Lehigh University, Spring 2016
2. Warren York Dissertation Fellowship, Lehigh University, Fall 2014
3. Merit-Based Scholarship, UIBE , 2006-2010

Affiliations

American Economic Association, Southern Economic Association, American Society of Health Economists, Chinese Economist Society.

Skills

Computer Software:
STATA, SAS, Mathmatica, LaTeX

Languages:
English (proficient) and Mandarin (native)

REFERENCES

Chad D. Meyerhoefer
Associate Professor of Economics,
Department of Economics
Lehigh University
Rauch Business Center,
621 Taylor Street,
Bethlehem, PA 18015
Email: chm308@lehigh.edu
(Tel) 610-758-3445

James A. Dearden
Professor and Chair of Economics,
Department of Economics
Lehigh University
Rauch Business Center,
621 Taylor Street,
Bethlehem, PA 18015
Email: jad8@lehigh.edu
(Tel) 610-758-5129

Stephen G. Buell
Professor of Finance,
Perella Department of Finance
Lehigh University
Rauch Business Center,
621 Taylor Street,
Bethlehem, PA 18015
Email: steve@lehigh.edu
(Tel) 610-758-3436