# THREE ESSAYS IN HEALTH AND LABOR ECONOMICS

by

BHAVNEET WALIA

B.A., PANJAB UNIVERSITY, 2001
M.A., PANJAB UNIVERSITY, 2003

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Economics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

# Abstract

The dissertation examines empirical issues related to health and labor economics. It has long been debated whether breastfeeding leads to a higher intelligence quotient (IQ) and greater scholastic achievement. The first study empirically examines the issue. Many past studies fail to take into account the possible endogeneity of the breastfeeding decision and thus falsely identify the correlation between breastfeeding and IQ as a causal relationship. We attempt to distinguish the causation and correlation between the two variables. Our results show that, after controlling for possible endogeneity, breastfeeding has no significant impact on IQ or scholastic achievement.

The second essay examines the link between breastfeeding and childhood obesity. Heath economics researchers view breastfeeding as a determining factor as to whether a child becomes obese. There are many theories, involving both biological and psychological factors, as to why breastfeeding is negatively linked to childhood obesity. This essay argues that the breastfeeding decision is not an exogenous one, so estimation technique such as ordinary least squares is not the correct way to estimate the relationship between breastfeeding and childhood obesity. Instruments are used to generate exogenous variations in the breastfeeding variable. After correcting for any estimation bias due to the breastfeeding variable being endogenous, this study documents the benefits of breastfeeding.

The third essay analyzes 19 semesters of student evaluations at Kansas State University. Faculty fixed effects are sizable and indicate that, as assessed by students, the best principles teachers also tend to be the best non-principles teachers. OLS estimates are biased because principles teachers are drawn from the top of the distribution and because unmeasured faculty characteristics are correlated with such variables as the response rate and student effort. Student ratings are lowest for new faculty but stabilize quickly. Expected GPA of the class is not an important determinant of student ratings, but equitable grading is; and the rewards for equitable grading appear larger for principles classes. The lower ratings in principles classes are fully accounted for by greater class size.

# THREE ESSAYS IN HEALTH AND LABOR ECONOMICS

by

BHAVNEET WALIA

B.A., PANJAB UNIVERSITY, 2001
M.A., PANJAB UNIVERSITY, 2003

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Economics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2008

Approved by:

Major Professor
Dr. Dong Li

# Abstract

The dissertation examines empirical issues related to health and labor economics. It has long been debated whether breastfeeding leads to a higher intelligence quotient (IQ) and greater scholastic achievement. The first study empirically examines the issue. Many past studies fail to take into account the possible endogeneity of the breastfeeding decision and thus falsely identify the correlation between breastfeeding and IQ as a causal relationship. We attempt to distinguish the causation and correlation between the two variables. Our results show that, after controlling for possible endogeneity, breastfeeding has no significant impact on IQ or scholastic achievement.

The second essay examines the link between breastfeeding and childhood obesity. Heath economics researchers view breastfeeding as a determining factor as to whether a child becomes obese. There are many theories, involving both biological and psychological factors, as to why breastfeeding is negatively linked to childhood obesity. This essay argues that the breastfeeding decision is not an exogenous one, so estimation technique such as ordinary least squares is not the correct way to estimate the relationship between breastfeeding and childhood obesity. Instruments are used to generate exogenous variations in the breastfeeding variable. After correcting for any estimation bias due to the breastfeeding variable being endogenous, this study documents the benefits of breastfeeding.

The third essay analyzes 19 semesters of student evaluations at Kansas State University. Faculty fixed effects are sizable and indicate that, as assessed by students, the best principles teachers also tend to be the best non-principles teachers. OLS estimates are biased because principles teachers are drawn from the top of the distribution and because unmeasured faculty characteristics are correlated with such variables as the response rate and student effort. Student ratings are lowest for new faculty but stabilize quickly. Expected GPA of the class is not an important determinant of student ratings, but equitable grading is; and the rewards for equitable grading appear larger for principles classes. The lower ratings in principles classes are fully accounted for by greater class size.

# Table of Contents

# List of Figures

# List of Tables

Acknowledgements

Thank you to everyone who has helped me toward the completion of this dissertation. Specifically, I would like to thank my parents, Mr. Mohan Deep Singh and Mrs. Savinder Kaur, who encouraged me to pursue higher studies in economics. Thanks to Harkamal Walia for his daily inspiration and guidance. Thank you to Jeewan Jyot for making a special trip to help me through my qualifying exam preparation.

Thank you to Dr. Dong Li for helping me build my research skills from zero. I feel that I can build upon this foundation for years to come. Thank you to Dr. Ragan for showing great confidence in me throughout our collaborations. My deep interest in labor economics developed directly from your courses. Thank you to Dr. Chang for showing me how to conduct theoretical research and for making my first publication possible during graduate school. Thank you to Dr. Thomas and Dr. Blankenau for helping me develop my teaching ability and for their useful advice during my job market preparation. Thank you to Dr. Boyer and Dr. Bachmeier for helpful inputs and advice toward the completion of my dissertation. A special thanks to Dr. Hancock for her support and encouragement.

I would like to thank Shane Sanders for always being a great friend and for always supporting my academic efforts. I would also like to recognize the great friends I have made in graduate school specifically, Jasdeep Saini, Hana Janoudova, Amanda Freeman, Casey Abington, Yaseen Alhaj, Jaime Andersen, Dave Brown, Shin-Jae Kang, Eddery Lam, Canh Le, Andrew Ojede, Burak Onemli, Kyle Ross, Ruben Sargsyan, Renfeng Xiao, Joel Potter, Alexandra Gregory, Abhinav Alakshendra, Uma Sarmistha, Rashmi Dhankhar, Sandeep Rana and Kara Ross.

# Dedication

To my grandparents, Dr. Surjan Singh Alhuwalia, Mrs. Nirmal Kaur, Mr. Mehar Singh Bharara, and Mrs. Gurcharan Kaur Bharara, for showing me the value of learning.

# Preface

This dissertation explores issues in health and labor economics. The first two essays explore issues related to the academic and physical development of children. Specifically, they study the impact of breastfeeding on the academic achievement and physical development of children. Controlling for various other confounding factors, these two essays explore the causation between breastfeeding and childhood development. They correct the estimation bias that arises when the breastfeeding variable is treated as an exogenous decision. Depending on the main focus of the study, different sets of instruments are used to generate exogenous variations in the breastfeeding variable. Comparisons of results from ordinary least squares estimation and instrumental variable estimation (generalized methods of moment, two stage least squares & instrumental probit analysis) are presented in both the essays. The third essay shows an improvement on the ordinary least squares approach by using a fixed effect estimation technique that controls for individual characteristics of faculty members. It evaluates the differences in student evaluations of principles and non-principles classes and how the allocation of faculty occurs across these courses.

The first essay studies the impact of breastfeeding on childhood academic development. After controlling for maternal, child, and family characteristics, this essay examines the association between breastfeeding duration and academic achievement. It compares the results from ordinary least squares and generalized methods of moments estimation. Contrary to previous literature that treated the breastfeeding variable as exogenous, this research treats breastfeeding as a potentially endogenous variable. Instruments used to generate exogenous variations in the breastfeeding variable are maternal work hours, different types of child care arrangements when the child is one year old, and whether the child is raised in a single parent family. In this essay, I try to distinguish causation and correlation between breastfeeding and academic achievement. Treating an endogenous variable as exogenous could lead to overestimation of the impact of breastfeeding on academic achievement. Results from both ordinary least squares and generalized methods of moments approach are presented. After controlling for the breastfeeding variable endogenously, the significant impact of breastfeeding on various academic achievement tests found in ordinary least squares estimation disappears. This essay was jointly done with Dr. Dong Li.

The second essay studies physical aspects of childhood development. Specifically, it examines the impact of breastfeeding on childhood obesity. The breastfeeding decision is again treated as an endogenous variable. Most of the current research literature studying breastfeeding and childhood obesity has treated breastfeeding as an exogenous decision. This essay estimates the impact of breastfeeding on body mass index (BMI) of the child and on his/her probability of being obese. Two stage least squares and instrumental probit analyses are used to treat the breastfeeding variable as endogenous. Variables related to the number of educated adults in the household are employed to generate exogenous variations in the breastfeeding variable. Results from this study confirm the benefits of breastfeeding.

The third essay compares estimation results from ordinary least squares estimation and fixed effect estimation in the context of student evaluation of teaching. It looks at the differences in student evaluations of principles and non-principles classes and how the allocation of faculty occurs across these courses. Results from this study show that traditional OLS estimates are biased because principles teachers are drawn from the top of the distribution, and unmeasured faculty characteristics are correlated with such variables as the response rate. A fixed effect estimation technique is used to account for individual faculty characteristics. This essay was jointly done with Dr. Jim Ragan.

# CHAPTER 1 - Any Causal Link between Breastfeeding and Scholastic Achievement?

## I. Introduction

The impact of breastfeeding on the health and academic outcomes of children has been a longstanding issue of debate. In recent years, a large amount of literature has emerged showing the association between breastfeeding and cognitive development. A number of papers claim that breastfeeding has a positive effect on child cognitive development. At issue is whether such papers have properly distinguished between causation and correlation with respect to these two variables. In this study, we try to distinguish causation and correlation between duration of breastfeeding and indicators of cognitive development such as scholastic achievement and intelligence quotient (IQ).

Horwood and Fergusson (1998) examined whether breastfeeding causes cognitive benefits into "young adulthood." Controlling for factors such as mother's age, mother's education, family socio-economic status, average income of the family, average standard of living of the family, mother's smoking habits, number of siblings, and birth weight, the authors found breastfeeding to have a lasting positive effect on IQ and other academic outcomes. This study was highly publicized by CNN Headline News, among other news outlets, in January 1998 following its publication in *Pediatrics*. Children in the sample were tested on various standard scales such as the Wechsler Intelligence Scale for children (WISC-R) and the Progressive Achievement Test of Reading Comprehension (PAT). Out of ten different measures of academic achievement, the authors found breastfeeding to be positive and significant in nine cases. The only measure without a significant effect on breastfeeding is the teacher's rating of reading ability at age eight years. We are able to replicate their results using the same data set and the same linear regression models.

Mortensen et al. (2002) examined the relationship between duration of breastfeeding and adult intelligence using two different samples. They found a strong and positive association between duration of breastfeeding and adult intelligence. The authors controlled for family characteristics such as marital status, education, age, height, smoking habits of the mother, social status of the family, number of prior pregnancies for the mother, gestation age, birth weight, birth length, and delivery conditions. They found a positive association between duration of breastfeeding and both parental social status and education. The authors provided three possible reasons that breastfeeding and cognitive development may be positively correlated. First, the composition of human milk and that of infant formula may be different. Second, this correlation may reflect differences in the surroundings of the child, the mother-child interaction, and the mother's attitude towards the child. Third, some unidentified factors may be correlated with both the infant feeding methods and the outcome variables.

Michaelson et al. (2003) found positive effects of breastfeeding on cognitive and visual acuity of brain development. The authors suggested that this positive association may be due to reasons similar to those outlined in Mortensen et al. (2002). There have been some past studies questioning the relationship between breastfeeding and the cognitive growth of children. Jacobson et al. (1999) challenged the past literature associating breastfeeding directly with a higher intelligence quotient. They found that, after taking into consideration the mother's IQ and other parental factors, the effect of breastfeeding on the child's IQ disappears. Their results showed that the positive association between breastfeeding and IQ is due to genetic and socio-environmental factors. Angelsen et al. (2001) examined the effects of breastfeeding on cognitive development at age one and five years. They compared children who were breastfed for three to six months with those who were breastfed for less than three months. They found a positive effect of breastfeeding on mental development and a comparatively smaller yet still positive effect on motor development for children age 13 months to five years. The importance of controlling for parental education and maternal IQ can be seen in papers such as Malloy and Berendes (1998) and Jacobson et al. (1999). Malloy and Berendes (1998) examined the effect of breastfeeding on intellectual development using a sample from a relatively homogeneous population to see if previous results had been contaminated by differences in background. Their study compared children who were fed on formula milk with children who were breastfed. Although failing to include information on maternal and paternal education, they found breastfeeding to be significant in a linear regression. However, after controlling for maternal and paternal education, breastfeeding became insignificant. Assuming the decision of breastfeeding to be exogenous, these studies have used linear models to estimate the effect of breastfeeding on the cognitive development of the child.

The goal of this study is to examine, as an empirical exercise, whether treating breastfeeding as exogenous leads to upward biased results. Past studies have used ordinary linear squares (OLS) estimation or an equivalent methodology such as analysis of variance (ANOVA) to measure the effect of breastfeeding on IQ. Treating the decision to breastfeed as exogenous is questionable, as suggested by Mortensen et al. (2002) and Michaelson et al. (2003). If there is non-zero correlation between breastfeeding and the error term due to omitted variables, the OLS approach would bias the coefficient estimates. For example, if higher income leads to higher IQ and longer breastfeeding duration, omitting income in the IQ equation would incorrectly attributes the income effect to breastfeeding. An instrumental variable approach, which relies on variation in breastfeeding that does not correlate with the error term in the IQ equation, can overcome this problem. In this essay, we take into account the possible endogeneity of breastfeeding in our model. We examine whether breastfeeding is causing better academic achievement or whether breastfeeding is merely correlated with better academic achievement. Using generalized methods of moment estimation we look at the impact of breastfeeding on scholastic achievement. Some previous studies examined the health benefits of breastfeeding on child outcome while treating breastfeeding as endogenous. Senauer and Kassouf (2000) used a sample of Brazilian children to look at the health benefits of breastfeeding and the demand for medical assistance by children. Even after controlling for possible endogeneity of breastfeeding, they still found breastfeeding to have a significantly positive effect on the health outcome of the children. Barrera (1991) and the Cebu Study Team (1992) also estimated the child health production function while allowing breastfeeding to be endogenous. Barrera (1991) looked at the relationship between duration of unsupplemented breastfeeding and height. In the same study the author showed how the use of OLS estimation by treating breastfeeding as exogenous can lead to biased results. Correcting for heterogeneity and endogeneity in previous models, the Cebu Study Team (1992) found that the environment of the child, including household and community, impact the child's health production function.

Section II describes the data set used for this project. Section III presents the model and the estimation techniques. Section IV shows the results from generalized method of moments. Concluding remarks are presented in Section V.

## II. Data Set

The data set was graciously provided to us by L. John Horwood, who used the same data in Horwood and Fergusson (1998). The data set was collected as a part of the Christchurch Health and Development Study (CHDS). It studied 1,265 children over a period of 18 years. All children born between April 15, 1977 and August 5, 1977 in the urban region of Christchurch, New Zealand were included in the study. The children were studied at birth, at age one year, and then yearly until age 16. The final survey interview was conducted at age 18. Parents were concurrently surveyed to get information on family characteristics such as family income, standard of living, maternal and paternal characteristics, family size, and birth weight of child. See Horwood and Fergusson (1998) for more information on the data.

Some of the main covariates used in similar studies in the past include mother's age, maternal education, socio-economic status, family income, number of children in the family, standard of living, gender, birth weight, and mother's smoking habits. CHDS includes years of maternal education, ranging from no formal education to college graduate. Molly and Berendes (1998) considered mother's education in their study. Another potentially relevant variable, if measured correctly, is maternal IQ. CHDS does provide information on maternal IQ, but 602 of 1,265 observations are missing for this variable. The authors of the data indicate that such a large proportion of missing values is likely due to selection bias. The authors explain that those failing to provide this information were more likely to have poorer education. Therefore, we do not include this variable in our study.[1]

Family socio-economic status is based on the Elley/Irving scale. Father's occupation is a variable with information on whether the father is an unskilled, skilled, or professional worker.

---

[1] We imputed the missing values of this variable and added it into the main model. This did not change the results. Breastfeeding significantly affected IQ in the linear model and was insignificant in the GMM estimation.

Average family income represents the family's average gross income decile, with respect to all other sampled families, over each of the child's first five years. Past studies have shown that children from high income families tend to do better in academics due to the stability in their environment. A positive and healthy environment has a significant positive impact on a child's future outcomes as shown by the past literature.

Instrumental variables are used to generate exogenous variation in the breastfeeding variable. Instruments include maternal employment hours when the baby was four months old, the baby's age when the mother returned to work, number of hours the baby received care from extended family members (such as grandparents) or non-family members (such as daycares) at age one year, and whether the child was raised in a single parent family. Description and summary statistics for all the variables used in the study are presented in Table 1.1 and Table 1.2, respectively.

Ten standardized tests are used to calculate academic outcome of the child. The revised Wechsler Intelligence Scale for children is used to test the child's IQ level. This test reports the child's IQ at age eight years and nine years. Progressive Achievement Test (PAT) at age 10 years and 12 years is used to measure the reading comprehension abilities of the child. Teacher ratings are used to measure ability in reading and math at age eight and 12 years. A scholastic test for which scores fall between 0 and 69 is used to measure the aptitude for high school success. Figure 1.1 presents the graphs showing the relationship between the unadjusted test scores and duration of breastfeeding using cubic splines (Hastie and Tibshirani, 1990; Schimek, 2000).

## III. Model

The basic model used in this study is the same as the one used by Horwood and Fergusson (1998):

$$AA = \alpha + X\beta + BF\gamma + e$$

where AA stands for academic achievement (one of the ten standardized test scores), X includes family and child characteristics such as mother's age,[2] mother's education, family's socio-economic status, average family income, average standard of living of the family, mother's smoking habits, family size, and birth weight of the baby, and BF stands for duration of breastfeeding. We were able to replicate the OLS results in Horwood and Fergusson (1998). However, we argue that the OLS results are biased because the possible endogeneity is not properly accounted for.

We run generalized method of moments (GMM) regressions using age of the child when the mother returns to work, hours worked by the mother when the child was four months old, hours the child was under care by extended family members and non-family members per week, and if the family was a single parent or two-parent family as instruments. Our first two instruments are age of child when the mother returned to work (Mother Returns to Work) and the number of hours the mother worked when the baby was four months old (Maternal Work Hours). Longer hours of work would reduce the amount of time the mother spends with the baby. Using a sample of 668 black and 511 white women, Kurinu et al. (1989) found a negative relationship between time at which the mother returned to the workforce and duration of breastfeeding, while the results differed in magnitude for white and black women, full time and part time workers,

[2] We tried using the quadratic form of the maternal age variable to see if it significantly affected the academic outcome of the child. We did not find any significant impact of mother's age squared.

and between skilled and unskilled workers. The second and third instruments are the number of hours per week the baby is under care by extended family members such as grandparents and relatives (Family Care Hours), and the number of hours per week that the child was being cared for by non-relatives such as daycares (Other Care Hours). These instruments follow from the past literature that has showed the importance of family circumstances on the decision of breastfeeding. A study by Sullivan et al. (2004) showed how the load of household tasks could lead to an early cessation of breastfeeding. Duration of breastfeeding can be longer if the mother receives some direct or indirect help from other family members for taking care of the baby.[3] First stage results show a negative correlation between Family Care Hours and duration of breastfeeding, which opposes our a priori expectation. First stage results of this study show a negative relation between these two variables, which opposite to what we expected. The final instrument used was whether the family was a single parent family (Single Parent Family). Validity of this instrument comes from past literature, which shows that having family support and help in the household is a positive encouragement for the mother to breastfeed her child (Sullivan et al., 2004). By allowing the decision to breastfeed to be endogenous, we found that the significance of the breastfeeding coefficient is reduced and in some cases disappears. In other words, recognizing the endogeneity of breastfeeding within the model weakens any evidence that breastfeeding has a significant effect upon child academic outcome. There are two requirements of good instruments- relevance and exogeneity. To show the relevance of the instruments outlined above, first stage results are presented in Tables 1.3 (a) to (d). Relevance of the instruments is displayed through the significance of instruments in the first-stage

[3] Mother's age, maternal education, socioeconomic status of the family, average income of the family, maternal smoking, birth weight, and family size were found to be significant factors that increase the probability of longer duration of breastfeeding. The probit estimation technique was used to estimate these effects.

regression. The coefficient on Family Care Hours is significant in four of the ten models, where each model uses a different measure of academic achievement. The coefficient on Other Care Hours is significant at the .05 level in eight of the ten tests and at the .10 level in nine of the ten tests. Lastly, the coefficient on Maternal Work Hours is significant at the .10 level in one of the ten tests. Exogeneity of the instruments is displayed through insignificant Sargan statistic values in Tables 1.4(a) through 1.4(d).

## IV. Results

Tables 1.4 (a) to (d) present the results from OLS and GMM. We were able to replicate the results in Horwood and Fergusson (1998) based on OLS. For nine out of the 10 dependent variables breastfeeding is significant and positive in the OLS regressions. This lends support to the claim that breastfeeding is beneficial to academic achievement. However, when we use the GMM with the instruments discussed in the previous section, the significance of the breastfeeding effect largely declines.. Breastfeeding is insignificant at the 5% level for 9 out of the 10 models in the GMM regressions. Breastfeeding becomes insignificant for most of the tests when instruments such as mother's work hours and family environment are taken into account. We argue that the positive relationship between breastfeeding and academic achievement is merely an association. Once we control for the possible endogeneity of breastfeeding decision, breastfeeding is no longer significant. We were limited by the data to a small set of usable instrumental variables. Because of this restriction, the results from this research project are more suggestive than conclusive. The Sargan statistics[4] support the validity

[4] The Sargan statistic is used to test the exogeneity assumption of the instrumental variables. The null hypothesis is zero correlation between the instruments and the error term. Within the study, we cannot reject the null hypothesis of zero correlation. This result provides validation for our chosen instruments.

of the instruments used in this essay. Other variables, including family income, birth order, and birth weight, have significant effects on academic achievement, which is consistent with past literature.

Table 1.4(a) presents the results from the IQ tests and the reading test. The major change is in the coefficient of breastfeeding, which becomes insignificant after we control for endogeneity in the Horwood and Fergusson model. As we can see, mother's age has a significantly positive impact on IQ, as measured by the Revised Wechsler Intelligence scale. Using a sample of 11,742 siblings from the Netherlands, Kalmijn and Kraaykamp (2004) showed a positive association between maternal age and child's schooling after controlling for child surroundings. Angrist et al. (1996) found that children born to younger mothers have more difficulties in school. They are more likely to repeat classes during their academic career compared to children from older mothers. As shown in Tables 1.4 (a) to (d) maternal education has a positive effect on the academic outcome of the child. Educated mothers tend to be more stable and careful about the surroundings of their children. The environment in which a child grows up has a significant impact of his or her future outcome. This is consistent with the past literature. A positive home environment plays a significant positive role in the successful outcome of the children.

Average family income is a significant variable in OLS and GMM. The coefficients for this variable in the linear regression and GMM regression are very close to each other and significant at 5%. This positive coefficient is supported by Blanden et al. (2006). They showed that income has a positive impact on non-cognitive activities, which indirectly affect the academic outcome of the children. Plug and Wim (2005) examined adopted children to correct for the possible selection bias for high income parents having a genetic effect on their children, which would lead to better academic achievement. Their results supported the significantly positive relationship between family income and better academic outcomes. Similar studies have shown a significant long-term impact of family income level on cognitive and non-cognitive achievement of the child. The socio-economic status of the family has a positive impact on the academic outcome of the child. The information on this variable is collected using the Elley/Irving scale of socio-economic status, which looks at the father's occupation type, whether it is managerial, skilled, or unskilled.

The significant effect of birth order can be seen in the results. Being born later among siblings has a negative impact on academic outcome of children. Booth and Hiau (2006) found that the birth order effects are persistent even after controlling for other family characteristics. Families who choose quantity sacrifice quality in terms of child academic achievement. Birth weight has a positive impact on the academic outcome of the child. The small significant effect of birth weight in our results is in line with the study by Miller et al. (2005).

Our results show that breastfeeding is not a significant factor in the child's academic outcomes once we properly take into account other factors that indirectly affect the decision to breastfeed in the first place. We find that the significance of breastfeeding in OLS could merely be a correlation. These results suggest that breastfeeding could be an endogenous variable, and treating it as an exogenous variable could lead to biased estimates.[5]

After controlling for possible endogeneity of breastfeeding, we find that breastfeeding has no significant impact on IQ or scholastic achievement. Mothers who cannot breastfeed their children due to certain reasons such as health problems or work constraints need not feel guilty.

[5] Results from the GMM model do show some signs of weak instruments (i.e., large increase in standard error of the coefficients in the GMM model compared to OLS). Given the availability of variables that could be used as instruments for breastfeeding, it was difficult to find a strong set of instruments.

Our study does not question or look into the significance of breastfeeding on health outcomes of the child, as established by past research. One main limitation of our study is that due to severe and non-random missing data we are unable to control for maternal IQ, which may have a significant impact on the academic outcome of the child. Our results show that the main confounding factors for the academic outcome of the child are maternal age, maternal education, family income, family size, gender, and birth weight.

## V. Conclusion

Some of the past research suggests a significantly positive association between breastfeeding and the child's academic outcomes. In this essay we argue that this result is based on the improper assumption that the breastfeeding decision is exogenous. We find that after controlling for the endogeneity of breastfeeding, the significance of breastfeeding found in past literature disappears. We speculate that it may be the amount of time a mother spends with her baby that causes higher IQ or better academic success for the child. Though we are unable to directly control for the amount of time a mother spends with the baby given the data set, its proxy, maternal employment, is supported by past literature. The results of this study could have an important implication for mothers who are unable to breastfeed their babies due to health or employment reasons. They should not feel guilty about such constraints. They can spend a good amount of time with their children and ensure the children's future academic outcomes regardless of their breastfeeding status. As established by past research, the environment in which the child grows up is very important for the successful outcome of the child.

Our study adds to the past literature by showing the importance of mother-child interactions. The main contribution of this essay is that breastfeeding should be considered an endogenous variable in the academic outcome regressions and OLS estimation can produce biased results when measuring the effect of breastfeeding on academic achievement of the child. Once the possible endogeneity of breastfeeding is properly taken into account, the significantly positive impact of breastfeeding on academic outcomes disappears. The weakness of available variables, in terms of generating exogenous variations in the breastfeeding variable, is one limitation of this study. Finding good instruments, ones that show strong correlation with breastfeeding and are not related to academic achievement, could provide further insights into this topic. Given the variables in the current data set, maternal work history, childcare arrangement, and family environment were the best instrumental variables available. Future work on finding stronger instruments for the breastfeeding variable can further improve the literature.

# VI. Figures and Tables

## Table 1.1 Description of CHDS Variables

| | |
|---|---|
| Mothers age | Mothers age in years |
| Maternal education | Maternal education at the child's birth |
| Less High School | 0/1: No formal educational qualifications |
| High School | 0/1: Secondary (high school) qualifications |
| College | 0/1: Tertiary (college) qualifications |
| Average standard of living | Averaged interviewer rating of standard of living (5-point scale) |
| Family socio-economic status | Family socio-economic status at the time of the survey child's birth based on the Elley/Irving scale of socio-economic status for New Zealand applied to the father's occupation. Coded into three levels: |
| Professional | 0/1: Professional, Managerial |
| Clerical | 0/1: Clerical, Techincal, Skilled |
| Semiskilled | 0/1: Semiskilled, Unskilled, Unemployed |
| Birth Order | The number of children in the family at the child's birth.Range:1 – 5 |
| Birth Weight | The child's birth weight in grames |
| Average family income | Averaged family income decile.Range: 1.0-10.0 with lower scores implying lower incomes |
| Mothers smoking habits | Maternal smoking during pregnancy (cigarettes per day). Range: 0-50 |
| Maternal employment at age 4 months | Maternal workforce participation (hrs worked per week) when child aged 4 months (range: 0-98) |
| Childs age at mothers return to work | Child's age (months) when mother first re-entered the paid workforce after birth. A code of 99 implies that mother never re-entered workforce. A code of zero implies the child was aged <1 month when mother re-entered workforce. |
| Under care of family member | The number of hours per week that the child was being cared for by other family members (eg grandparents, older children, other relatives) at age 1 year. |
| Under care of non family member | The number of hours per week that the child was being cared for by non-relatives (eg day care centre, paid child minder, etc) at age 1 year. |
| Single parent family | 0/1: if it is a single parent family |
| Male | 0/1 : child's sex |
| Duration of Breastfeeding | Duration of breastfeeding, range 0-12 |

*Note: These variable definitions were provided by John Horwood.*

**Table 1.2 Summary Statistics of CHDS Variables**

| Variable Description<br>Family & child characteristics | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| Mothers age(in years) | 25.809 | 4.898 | 15 | 47 |
| Maternal education | | | | |
| Less High School | .511 | .500 | 0 | 1 |
| High School | .303 | .459 | 0 | 1 |
| College | .186 | .389 | 0 | 1 |
| Average standard of living | 2.893 | .462 | 1 | 5 |
| Family socio-economic status | 2.068 | .683 | 1 | 3 |
| Paternal education | | | | |
| Less High School | .483 | .499 | 0 | 1 |
| High School | .334 | .472 | 0 | 1 |
| College | .183 | .387 | 0 | 1 |
| Birth Order | 1.975 | 1.000 | 1 | 5 |
| Birth Weight (grams) | 3356.536 | 528.753 | 1100 | 5140 |
| Average family income | 5.682 | 2.482 | 1 | 10 |
| Mothers smoking habits(per day) | 4.105 | 7.820 | 0 | 50 |
| Maternal employment at age 1 year | 3.262 | 8.666 | 0 | 98 |
| Childs age at mothers return to work | 76.132 | 40.402 | 0 | 99 |
| Under care of family member | 1.550 | 5.997 | 0 | 98 |
| Under care of non family member | 1.159 | 5.125 | 0 | 98 |
| Gender | 1.498 | .500 | 1 | 2 |
| Duration of breastfeeding | 3.974 | 4.198 | 0 | 12 |

*Notes: Maternal education, family socio-economic status, paternal education, is coded on a 3 point scale. Childs birth weight is measures in grams. Maternal employment is measures in hours worked per week.*

**Table 1.2 (continued) Summary Statistics of CHDS Variables**

| Variable Description<br>Satndardized tests | Number of observations | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Revised Wechsler Intelligence Scale for IQ at age 8 years | 881 | 101.88 | 15.60 | 30 | 143 |
| Revised Wechsler Intelligence Scale for IQ at age 9 years | 811 | 104.05 | 16.82 | 40 | 150 |
| Teacher rating of reading ability at 8 years | 1081 | 2.45 | 1.09 | 1 | 5 |
| Teacher rating of reading ability at 12 years | 1006 | 2.34 | 1.08 | 1 | 5 |
| Teacher rating mathematics at 8 years | 1081 | 2.58 | 1.01 | 1 | 5 |
| Teacher rating mathematics at 12 years | 999 | 2.49 | 1.08 | 1 | 5 |
| Progressive Achievement test of reading comprehension at age 10 years | 847 | 10.38 | 7.06 | 0 | 31 |
| Progressive Achievement test of mathematics at age 11 years | 831 | 24.91 | 7.42 | 0 | 40 |
| Progressive Achievement test of reading comprehension at age 12 years | 804 | 12.92 | 4.79 | 0 | 22 |
| Test of Scholastic achievement at age 13 years | 784 | 34.69 | 15.12 | 0 | 69 |

*Note: Test of scholastic achievement is designed to measure the high school success of the child*

**Table 1.3 (a) First Stage Results**

| | | Revised Wechsler Intelligence Scale for IQ at age 8 years | Revised Wechsler Intelligence Scale for IQ at age 9 years | Teacher rating of reading ability at 8 years |
|---|---|---|---|---|
| | | **Breastfeeding** | **Breastfeeding** | **Breastfeeding** |
| Maternal age at the child's birth | | -0.013 | -0.028 | -0.009 |
| | | (0.372) | (0.768 | (0.278) |
| Maternal education at the child's birth; | High School | 1.169*** | 1.122*** | 1.317*** |
| | | (3.436) | (3.163) | (4.228) |
| | College | 3.171*** | 3.023*** | 3.321*** |
| | | (7.430) | (6.927) | (8.697) |
| Family socio-economic status; | Professional | 1.291*** | 1.378*** | 1.023** |
| | | (2.660) | (2.781) | (2.347) |
| | Clerical | 0.659** | 0.662** | 0.558** |
| | | (2.182) | (2.131) | (2.000) |
| Rating of standard of living | | 0.335 | -0.144 | 0.237 |
| | | (0.771) | (0.318) | (0.608) |
| Average family income | | -0.001 | -0.080 | -0.033 |
| | | (0.016) | (1.076) | (0.501) |
| Maternal smoking during pregnancy | | -0.079*** | -0.079*** | -0.078*** |
| | | (5.166) | (4.571) | (5.725) |
| Gender | | -0.097 | 0.014 | -0.322 |
| | | (0.369) | (0.052) | (1.341) |
| Birth order; | Second | -0.490 | -0.373 | -0.436 |
| | | (1.553) | (1.131) | (1.534) |
| | Third | 0.116 | 0.747 | 0.188 |
| | | (0.264) | (1.584) | (0.473) |
| | Fourth | -0.441 | -0.477 | -0.330 |
| | | (0.688) | (0.739) | (0.551) |
| | Fifth | 1.494 | 1.975* | 1.417 |
| | | (1.584) | (1.857) | (1.542) |
| Child's birth weight | | 0.001*** | 0.000* | 0.001*** |
| | | (2.786) | (1.901) | (3.284) |
| Family care hours | | -0.041* | -0.037* | -0.027 |
| | | (1.897) | (1.653) | (1.495) |
| Mother returns to work | | -0.002 | -0.003 | -0.000 |
| | | (0.585) | (0.860) | (0.131) |
| Maternal work hours | | -0.017 | -0.022 | -0.013 |
| | | (0.681) | (0.860) | (0.564) |
| Other care hours | | -0.027* | -0.040 | -0.031** |
| | | (1.655) | (1.586) | (1.968) |
| Single parent family | | -0.317 | -0.505 | -0.222 |
| | | (0.598) | (0.916) | (0.464) |
| Constant | | 0.412 | 3.132 | 0.615 |
| | | (0.196) | (1.418) | (0.325) |
| N | | 849 | 779 | 1041 |

Note:* indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01; t-statistics are reported in the parentheses;

**Table 1.3 (b) First Stage Results**

| | Teacher rating of reading ability at 12 years | Teacher rating mathematics at 8 years | Teacher rating mathematics at 12 years |
|---|---|---|---|
| | **Breastfeeding** | **Breastfeeding** | **Breastfeeding** |
| Maternal age at the child's birth | -0.025 | -0.009 | -0.023 |
| | (0.750) | (0.278) | (0.695) |
| Maternal education at the child's birth; High School | 1.315*** | 1.317*** | 1.291*** |
| | (4.001) | (4.228) | (3.919) |
| College | 3.157*** | 3.321*** | 3.041*** |
| | (8.030) | (8.697) | (7.734) |
| Family socio-economic status; Professional | 1.008** | 1.023** | 1.035** |
| | (2.242) | (2.347) | (2.295) |
| Clerical | 0.629** | 0.558** | 0.607** |
| | (2.149) | (2.000) | (2.059) |
| Rating of standard of living | 0.109 | 0.237 | 0.081 |
| | (0.254) | (0.608) | (0.189) |
| Average family income | -0.034 | -0.033 | -0.039 |
| | (0.487) | (0.501) | (0.565) |
| Maternal smoking during pregnancy | -0.074*** | -0.078*** | -0.076*** |
| | (4.907) | (5.725) | (5.037) |
| Gender | -0.346 | -0.322 | -0.334 |
| | (1.371) | (1.341) | (1.322) |
| Birth order; Second | -0.312 | -0.436 | -0.364 |
| | (1.040) | (1.534) | (1.211) |
| Third | 0.355 | 0.188 | 0.357 |
| | (0.841) | (0.473) | (0.836) |
| Fourth | -0.376 | -0.330 | -0.393 |
| | (0.582) | (0.551) | (0.610) |
| Fifth | 1.431 | 1.417 | 1.410 |
| | (1.488) | (1.542) | (1.470) |
| Child's birth weight | 0.001*** | 0.001*** | 0.001*** |
| | (3.607) | (3.284) | (3.552) |
| Family care hours | -0.030 | -0.027 | -0.030 |
| | (1.561) | (1.495) | (1.574) |
| Mother returns to work | -0.002 | -0.000 | -0.002 |
| | (0.711) | (0.131) | (0.660) |
| Maternal work hours | -0.023 | -0.130 | -0.021 |
| | (0.879) | (0.564) | (0.789) |
| Other care hours | -0.037** | -0.031** | -0.038** |
| | (2.330) | (1.968) | (2.353) |
| Single parent family | -0.255 | -0.222 | -0.217 |
| | (0.491) | (0.464) | (0.415) |
| Constant | 1.148 | 0.615 | 1.276 |
| | (0.565) | (0.325) | (0.627) |
| N | 967 | 1041 | 961 |

Note: * indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01; t-statistics are reported in the parentheses

**Table 1.3 (c) First Stage Results**

| | | Progressive Achievement test of reading comprehension at age 10 years | Progressive Achievement test of mathematics at age 11 years |
|---|---|---|---|
| | | **Breastfeeding** | **Breastfeeding** |
| Maternal age at the child's birth | | -0.005 | -0.020 |
| | | (0.144) | (0.554) |
| Maternal education at the child's birth; | High School | 1.098*** | 1.225*** |
| | | (3.143) | (3.444) |
| | College | 3.204*** | 3.021*** |
| | | (7.372) | (6.883) |
| Family socio-economic status; | Professional | 1.089** | 1.203** |
| | | (2.224) | (2.383) |
| | Clerical | 0.592* | 0.666** |
| | | (1.922) | (2.117) |
| Rating of standard of living | | 0.227 | 0.109 |
| | | (0.501) | (0.236) |
| Average family income | | -0.004 | -0.012 |
| | | (0.057) | (0.161) |
| Maternal smoking during pregnancy | | -0.074*** | -0.074*** |
| | | (4.648) | (4.514) |
| Gender | | -0.070 | -0.117 |
| | | (0.260) | (0.426) |
| Birth order; | Second | -0.341 | -0.445 |
| | | (1.054) | (1.346) |
| | Third | 0.376 | 0.427 |
| | | (0.824) | (0.902) |
| | Fourth | -0.524 | -0.684 |
| | | (0.819) | (1.070) |
| | Fifth | 1.790* | 1.834* |
| | | (1.791) | (1.834) |
| Child's birth weight | | 0.001*** | 0.001*** |
| | | (3.140) | (2.752) |
| Family care hours | | -0.036 | -0.040* |
| | | (1.550) | (1.650) |
| Mother returns to work | | -0.004 | -0.004 |
| | | (1.052) | (1.127) |
| Maternal work hours | | -0.040 | -0.021 |
| | | (1.554) | (0.820) |
| Other care hours | | -0.040*** | -0.041** |
| | | (2.661) | (2.440) |
| Single parent family | | -0.472 | -0.408 |
| | | (0.893) | (0.723) |
| Constant | | 0.311 | 1.378 |
| | | (0.145) | (0.628) |
| N | | 812 | 797 |

Note:* indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01; t-statistics are reported in the parentheses

**Table 1.3 (d) First Stage Results**

| | Progressive Achievement test of reading comprehension at age 12 years | Test of Scholastic achievement at age 13 years |
|---|---|---|
| | **Breastfeeding** | **Breastfeeding** |
| Maternal age at the child's birth | -0.031 | -0.041 |
| | (0.843) | (1.122) |
| Maternal education at the child's birth; High School | 1.445*** | 1.422*** |
| | (3.984) | (3.898) |
| College | 3.091*** | 3.187*** |
| | (6.996) | (7.147) |
| Family socio-economic status; Professional | 1.168** | 1.229** |
| | (2.306) | (2.373) |
| Clerical | 0.639** | 0.600* |
| | (1.997) | (1.823) |
| Rating of standard of living | 0.103 | 0.173 |
| | (0.220) | (0.363) |
| Average family income | -0.019 | -0.024 |
| | (0.245) | (0.305) |
| Maternal smoking during pregnancy | -0.075*** | -0.073*** |
| | (4.549) | (4.352) |
| Gender | -0.037 | -0.041 |
| | (0.133) | (0.145) |
| Birth order; Second | -0.274 | -0.121 |
| | (0.816) | (0.351) |
| Third | 0.653 | 0.803* |
| | (1.379) | (1.673) |
| Fourth | -0.490 | -0.604 |
| | (0.737) | (0.931) |
| Fifth | 2.024** | 2.520** |
| | (2.069) | (2.387) |
| Child's birth weight | 0.001*** | 0.001*** |
| | (3.264) | (3.129) |
| Family care hours | -0.030 | -0.054* |
| | (1.257) | (1.879) |
| Mother returns to work | -0.004 | -0.005 |
| | (1.137) | (1.325) |
| Maternal work hours | -0.038 | -0.049* |
| | (1.389) | (1.699) |
| Other care hours | -0.038** | -0.046* |
| | (2.437) | (1.873) |
| Single parent family | -0.619 | -0.492 |
| | (1.117) | (0.818) |
| Constant | 1.144 | 1.326 |
| | (0.520) | (0.596) |
| N | 772 | 754 |

Note:* indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01;  t-statistics are reported in the parentheses

## Table 1.4 (a) OLS and GMM estimates

| Independent Variables | | Revised Wechsler Intelligence Scale for IQ at age 8 years | | Revised Wechsler Intelligence Scale for IQ at age 9 years | | Teacher rating of reading ability at 8 years | |
|---|---|---|---|---|---|---|---|
| | | OLS | IV | OLS | IV | OLS | IV |
| Duration of breastfeeding | | 0.243*** | 0.554 | 0.195** | 0.104 | 0.01 | 0.226* |
| | | (2.878) | (0.548) | (2.239) | (0.090) | (1.360) | (1.837) |
| Maternal age at the child's birth | | 0.227*** | 0.231*** | 0.235*** | 0.240** | 0.026*** | 0.028*** |
| | | (2.828) | (2.669) | (2.832) | (2.558) | (3.566) | (2.911) |
| Maternal education : | High School | 1.287 | 0.837 | 1.660** | 1.846 | 0.156** | -0.13 |
| | | (1.626) | (0.558) | (2.054) | (1.170) | (2.182) | (0.680) |
| | College | 3.451*** | 2.325 | 4.364*** | 4.656 | 0.284*** | -0.442 |
| | | (3.450) | (0.672) | (4.289) | (1.263) | (3.176) | (1.034) |
| Family socio-economic status: | Professional | 1.945* | 1.556 | 2.818** | 2.868 | 0.095 | -0.133 |
| | | (1.689) | (0.898) | (2.421) | (1.524) | (0.923) | (0.708) |
| | Clerical | 1.061 | 0.684 | 0.784 | 0.616 | 0.087 | -0.046 |
| | | (1.319) | (0.634) | (0.957) | (0.542) | (1.183) | (0.386) |
| Rating of standard of living | | 1.346 | -1.342 | -1.843* | -1.834* | -0.041 | -0.091 |
| | | (1.387) | (1.273) | (1.855) | (1.737) | (0.462) | (0.767) |
| Average family income | | 0.387** | 0.389** | 0.423** | 0.415** | 0.050*** | 0.056*** |
| | | (2.282) | (2.265) | (2.397) | (2.087) | (3.244) | (2.840) |
| Maternal smoking during pregnancy | | 0.003 | 0.017 | 0.019 | 0.012 | 0.001 | 0.018* |
| | | (0.066) | (0.187) | (0.384) | (0.108) | (0.2050 | (1.680) |
| Gender | | 0.342 | 0.324 | 0.966 | 0.938 | -0.418*** | -0.352*** |
| | | (0.532) | (0.502) | (1.468) | (1.456) | (7.209) | (4.108) |
| Birth order: | Second | -2.298*** | -2.143** | -2.189*** | -2.227*** | -0.265*** | -0.185* |
| | | (2.978) | (2.446) | (2.798) | (2.641) | (3.813) | (1.905) |
| | Third | -1.861* | -1.914* | -1.666* | -1.677 | -0.329*** | -0.381*** |
| | | (1.941) | (1.965) | (1.670) | (-1.276) | (3.753) | (3.192) |
| | Fourth | -3.770** | -3.418** | -3.420** | -3.346* | -0.324** | -0.253 |
| | | (2.310) | (2.121) | (2.059) | (1.930) | (2.221) | (1.284) |
| | Fifth | 3.235 | 3.797 | -3.818* | -4.093 | -0.762*** | -1.056*** |
| | | (1.558) | (1.602) | (1.706) | (1.363) | (3.757) | (3.274) |
| Child's birth weight | | 0.002*** | 0.002** | 0.001** | 0.001* | 0.000*** | 0 |
| | | (3.442) | (2.117) | (2.171) | (1.674) | (4.125) | (0.801) |
| Sargan statistic | | | 4.811 | | 3.382 | | 0.596 |
| P-value | | | 0.307 | | 0.496 | | 0.963 |

Note: * indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01; t-statistics are reported in the parentheses

## Table 1.4 (b) OLS and GMM estimates

| Independent Variables | | Teacher rating of reading ability at 12 years | | Teacher rating of math at 8 years | | Teacher rating of math at 12 years | |
|---|---|---|---|---|---|---|---|
| | | OLS | IV | OLS | IV | OLS | IV |
| Duration of breastfeeding | | 0.020** | 0.131 | 0.017** | 0.083 | 0.024*** | 0.111 |
| | | (2.561) | (1.287) | (2.140) | (0.784) | (3.033) | (1.018) |
| Maternal age at the child's birth | | 0.015** | 0.019** | 0.019** | 0.020** | 0.020*** | 0.022** |
| | | (2.025) | (2.220) | (2.478) | (2.519) | (2.588) | (2.742) |
| Maternal education : High School | | 0.159** | 0.021 | 0.140* | 0.058 | 0.197*** | 0.097 |
| | | (2.159) | (0.132) | (1.887) | (0.353) | (2.616) | (0.573) |
| College | | 0.434*** | 0.094 | 0.220** | -0.001 | 0.303*** | 0.057 |
| | | (4.768) | (0.277) | (2.366) | (0.003) | (3.267) | (0.163) |
| Family socio-economic status: Professional | | 0.266** | 0.15 | 0.029 | -0.04 | 0.122 | 0.031 |
| | | (2.533) | (1.970) | (0.272) | (0.249) | (1.137) | (0.192) |
| | Clerical | 0.119 | 0.043 | 0.047 | 0.005 | 0.006 | -0.06 |
| | | (1.575) | (0.392) | (0.618) | (0.049) | (0.080) | (0.538) |
| Rating of standard of living | | -0.034 | -0.023 | -0.106 | -0.105 | 0.021 | 0.038 |
| | | (0.370) | (0.218) | (1.160) | (1.055) | (0.225) | (0.360) |
| Average family income | | 0.047*** | 0.051*** | 0.050*** | 0.052*** | 0.067*** | 0.070*** |
| | | (2.997) | (2.935) | (3.142) | (3.219) | (4.129) | (3.964) |
| Maternal smoking during pregnancy | | -0.002 | 0.007 | -0.001 | 0.005 | -0.002 | 0.005 |
| | | (0.357) | (0.812) | (0.142) | (0.559) | (0.529) | (0.570) |
| Gender | | -0.308*** | -0.278*** | -0.156*** | -0.137* | -0.188*** | -0.160** |
| | | (5.179) | (3.791) | (2.585) | (1.898) | (3.090) | (2.141) |
| Birth order: | Second | -0.201*** | -0.165** | -0.091 | -0.063 | -0.159** | -0.121 |
| | | (2.818) | (2.099) | (1.265) | (0.753) | (2.182) | (1.447) |
| | Third | -0.264*** | -0.311*** | -0.131 | -0.147 | -0.211** | -0.233** |
| | | (2.914) | (2.938) | (1.445) | (1.458) | (2.272) | (2.104) |
| | Fourth | -0.541*** | -0.487*** | -0.198 | -0.175 | -0.365** | -0.305* |
| | | (3.569) | (2.749) | (1.306) | (1.112) | (2.364) | (1.627) |
| | Fifth | -0.256 | -0.406* | -0.356* | -0.460* | -0.403* | -0.489* |
| | | (1.210) | (1.798) | (1.692) | (1.891) | (1.874) | (1.840) |
| Child's birth weight | | 0.000** | 0.000 | 0.000*** | 0.000 | 0.000** | 0.000 |
| | | (2.162) | (0.489) | (3.107) | (1.532) | (2.174) | (0.603) |
| Sargan statistic | | | 4.925 | | 1.343 | | 3.062 |
| P-value | | | 0.295 | | 0.854 | | 0.547 |

Note: * indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01; t-statistics are reported in the parentheses

## Table 1.4 (c) OLS and GMM estimates

| Independent Variables | | Progressive Achievement test of reading comp. at 10 years | | Progressive Achievement test of math at age 11 years | |
|---|---|---|---|---|---|
| | | OLS | IV | OLS | IV |
| Duration of breastfeeding | | 0.239*** | 0.891 | 0.211** | 1.743 |
| | | (2.762) | (1.171) | (2.446) | (1.494) |
| Maternal age at the child's birth | | 0.162** | 0.164* | 0.181** | 0.206** |
| | | (1.982) | (1.888) | (2.193) | (2.001) |
| Maternal education : High School | | 1.363* | 0.633 | 0.215 | -1.624 |
| | | (1.690) | (0.528) | (0.263) | (0.894) |
| College | | 4.204*** | 2.086 | 2.774*** | -1.883 |
| | | (4.091) | (0.784) | (2.689) | (0.501) |
| Family socio-economic status: Professional | | 2.221* | 1.485 | 2.658** | 0.756 |
| | | (1.909) | (1.017) | (2.232) | (0.379) |
| | Clerical | 1.210 | 0.763 | 1.305 | 0.170 |
| | | (1.478) | (0.771) | (1.567) | (0.130) |
| Rating of standard of living | | -1.654* | -1.792* | -1.344 | -1.341 |
| | | (1.656) | (1.737) | (1.335) | (1.088) |
| Average family income | | 0.319* | 0.318* | 0.598*** | 0.589*** |
| | | (1.847) | (1.834) | (3.406) | (2.920) |
| Maternal smoking during pregnancy | | -0.002 | 0.049 | 0.012 | 0.13 |
| | | (0.035) | (0.686) | (0.252) | (1.208) |
| Gender | | -2.535*** | -2.783*** | -1.193* | -1.061 |
| | | (3.868) | (2.463) | (1.800) | (1.331) |
| Birth order: | Second | -2.981*** | -2.831*** | -2.368*** | -1.845* |
| | | (3.809) | (3.418) | (2.985) | (1.851) |
| | Third | -2.459** | -2.743** | -1.296 | -2.042 |
| | | (2.490) | (2.463) | (1.286) | (1.593) |
| | Fourth | -2.15 | -1.859 | -3.198* | -2.041 |
| | | (1.280) | (0.094) | (1.894) | (0.964) |
| | Fifth | -4.156* | -5.179** | -4.235** | -6.966** |
| | | (1.947) | (2.299) | (1.972) | (2.449) |
| Child's birth weight | | 0.001** | 0.001 | 0.001** | 0.001 |
| | | (2.022) | (0.954) | (2.240) | (0.485) |
| Sargan statistic | | | 0.179 | | 1.783 |
| P-value | | | 0.996 | | 0.775 |

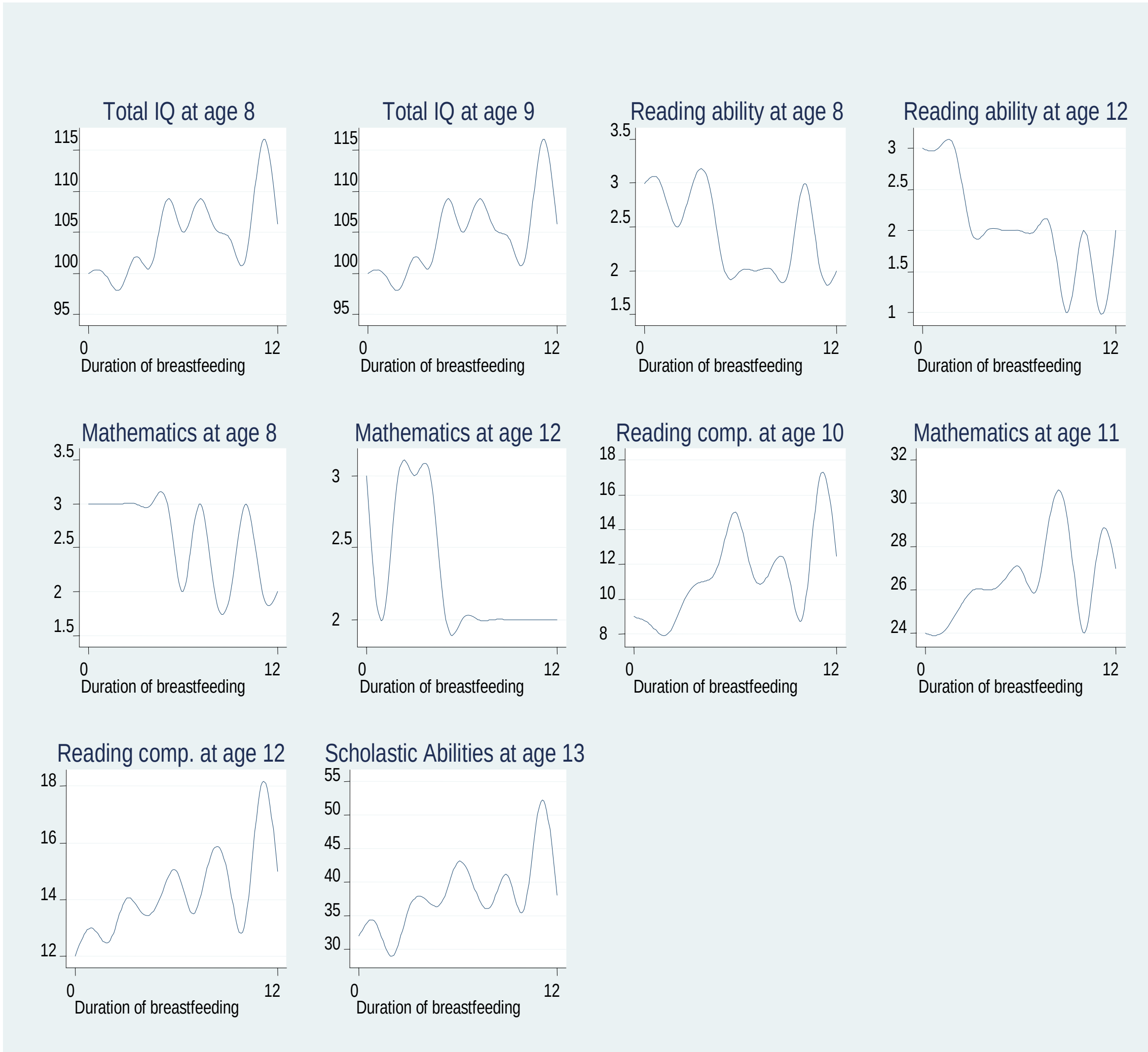
Note: * indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01; t-statistics are reported in the parentheses

**Table 1.4 (d) OLS and GMM estimates**

| Independent Variables | | Progressive Achievement test of reading comp. at 12 years | | Test of Scholastic Achievement at age 13 years | |
|---|---|---|---|---|---|
| | | OLS | IV | OLS | IV |
| Duration of breastfeeding | | 0.236*** | 1.702 | 0.207** | 0.697 |
| | | (2.625) | (1.624) | (2.426) | (1.056) |
| Maternal age at the child's birth | | 0.116 | 0.156 | 0.222*** | 0.237*** |
| | | (1.370) | (1.469) | (2.715) | (2.530) |
| Maternal education : High School | | 1.15 | -0.949 | 1.102 | 0.544 |
| | | (1.352) | (0.501) | (1.357) | (0.424) |
| College | | 4.249*** | -0.299 | 4.299*** | 2.794 |
| | | (4.008) | (0.086) | (4.261) | (1.192) |
| Family socio-economic status: Professional | | 3.308*** | 1.583 | 3.227*** | 2.59 |
| | | (2.705) | (0.848) | (2.747) | (1.806) |
| Clerical | | 1.545* | 0.491 | 1.24 | 0.843 |
| | | (1.804) | (0.383) | (1.504) | (0.842) |
| Rating of standard of living | | -0.828 | -0.903 | -0.057 | 0.02 |
| | | (0.791) | (0.713) | (0.057) | (0.018) |
| Average family income | | 0.222 | 0.236 | 0.696*** | 0.708** |
| | | (1.220) | (1.124) | (3.998) | (3.910) |
| Maternal smoking during pregnancy | | -0.005 | 0.113 | -0.01 | 0.033 |
| | | (0.110) | (1.154) | (0.209) | (0.468) |
| Gender | | -1.580** | -1.575* | -2.281*** | -2.241*** |
| | | (2.306) | (2.017) | (3.481) | (3.310) |
| Birth order: | Second | -2.181*** | -2.054** | -3.237*** | -3.252*** |
| | | (2.655) | (2.150) | (4.123) | (3.992) |
| | Third | -1.858* | -2.997** | -3.232*** | -3.744*** |
| | | (1.799) | (2.221) | (3.266) | (3.158) |
| | Fourth | -3.458** | -2.725 | -3.399** | -3.017 |
| | | (1.965) | (1.156) | (1.989) | (1.554) |
| | Fifth | -2.591 | -5.327 | -6.154*** | -7.603** |
| | | (1.149) | (1.679) | (2.707) | (2.982) |
| Child's birth weight | | 0.002** | 0.001 | 0.001** | 0.001 |
| | | (2.448) | (0.475) | (2.330) | (1.520) |
| Sargan statistic | | | 2.344 | | 3.195 |
| P-value | | | 0.672 | | 0.526 |

Note: * indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01; t-statistics are reported in the parentheses

**Figure 1.1 Unadjusted Relationship between Breastfeeding and Scholastic Achievement**



# CHAPTER 2 - Link between Breastfeeding and Childhood Obesity

## I. Introduction

The increasing incidence of childhood obesity has garnered significant attention from researchers, policy makers, and concerned parents. Obesity is an energy imbalance problem. There are genetic factors that play a role in the onset of obesity, but most of the problems are attributed to excess energy (in the form of calories) consumed and too little energy expended(Paracchini, Pedotti, and Taioli, 2005). Childhood obesity is a particularly worrisome condition because it is the main cause for the increase in adult health conditions occurring in children as they grow. Diabetes, high blood pressure, and high cholesterol were once conditions occurring mainly in the adult population. Recently, however, overweight and obese children have become afflicted with these conditions. The most important issue related to childhood obesity research is to identify all significant factors related to the childhood obesity epidemic. This paper is unique from medical literature on the same subject in that it employs an econometric methodology. Use of instrumental variable regression, in which the direction of relationships can potentially be sorted out, has not been utilized by medical researchers studying childhood obesity. Further, optimal maternal leave policies depend upon an understanding as to the true benefits of breastfeeding. In this sense, the relationship between breastfeeding and obesity has important microeconomic policy implications.

Medical researchers view shorter duration of breastfeeding as a determining factor as to whether a child becomes obese. There are many theories, involving both biological and psychological factors, as to why breastfeeding is negatively linked to obesity. Breast milk has been discovered to contain leptin, which is an amino acid associated with body weight regulation. Leptin works by sending signals to the brain to induce a feeling of satiety (i.e., feeling full) and to promote energy expenditure (Paracchini, Pedotti, and Taioli, 2005). Whether breastfeeding reduces childhood obesity is not clear from the current literature. Strawn and Mei (2004) studied the protective relationship between pediatric obesity and breastfeeding by race. They found a significant negative relation between breastfeeding and weight for non-hispanic white children in the sample. The breastfeeding variable used in the Strawn and Mei study is a continuous variable from Pediatric Nutrition Surveillance System, with 12,587 mother-child pairs. They used interaction terms between breastfeeding and gender, race, mother's age, maternal education, and maternal age. Their results showed breastfeeding's protection against the child being underweight and overweight.[6] In a study that potentially complicates the established relationship between breastfeeding and childhood obesity, Li et al. (2003) showed increasing trends in childhood obesity and incidence of breastfeeding in Britain and the United States. This study is critical of previous studies concluding that breastfeeding reduces the incidence of childhood obesity. In past studies, the breastfeeding decision is modeled as an exogenous variable. Current research argues that the breastfeeding decision is not an exogenous one, and estimation techniques such as ordinary least squares are not the correct way to estimate the relationship between breastfeeding and childhood obesity. Instruments are used to generate exogenous variations in the breastfeeding variable. Thus, econometric methodologies can potentially enlighten this literature.

[6] CDC defines an adult as overweight if his or her BMI is between 25 and 29.9 and obese if the BMI is greater than or equal to 30. For children and teens BMI for age is used to calculate if the child is underweight or overweight. Given the age and sex of the child, if the child's BMI falls in less than the 5th percentile, the child is declared underweight and if the BMI is equal to or greater than the 95th percentile, the child is considered overweight.

According to the Center for Disease Control (CDC), the percentage of obese children in all age groups dramatically increased between 1980 to 2004. The percentage children aged two to five who were deemed to be obese increased from 5.0% to 13.9%, while this percentage increased from 6.5% to 18.8% for six to 11 year olds and from 5.0% to 17.4% for 12 to 19 year olds.[7] These increased percentages confirm that childhood obesity is an epidemic. Among various causes and remedies related to childhood obesity, the effect of breastfeeding on childhood obesity has received particularly strong attention from researchers. The benefits of breastfeeding on various health outcomes of children are established in past research. However, the effect of breastfeeding on the weight of the child is not as well understood. Does breastfeeding provide protection against childhood obesity? In past studies, the breastfeeding decision has been modeled exogenously. This study uses the predicted breastfeeding (exogenous) variable to estimate the effect of breastfeeding on the Body Mass Index (BMI) of the child.

Hediger et al. (2001) found that maternal weight significantly affects obesity risk in a child, whereas breastfeeding does not directly reduce the prevalence of childhood obesity. Several authors have studied the dose-response relationship between breastfeeding duration and risk of childhood obesity. The commonly found negative relation between breastfeeding and prevalence of obesity could be due to hormonal factors in breast milk or certain maternal health factors. After controlling for various confounding factors such as maternal weight and maternal age at time of birth, authors do not find any dose-response relationship between the above-stated factors. Breastfed children were found to be at a lower risk of obesity, but the risk level did not

[7] These data are reported on the CDC website (http://www.cdc.gov/nccdphp/dnpa/obesity/childhood/prevalence.htm) and were generated using NHANES survey data from 1976-1980 and 2003-2004.

decline with longer breastfeeding duration. The issue of possible endogeneity of the breastfeeding decision was not addressed by the authors. We study the relationship between breastfeeding or duration of breastfeeding and risk of childhood obesity.

As Heitmueller (2007) treating an endogenous variable as exogenous could lead to biased estimation. He studied the impact of providing informal care on employment decisions. He found that care giving and labor force participation could be endogenous and not accounting for this endogeneity leads to an overestimate of the impact of care giving on labor force participation. Few childhood obesity studies control for the endogeneity of the breastfeeding decision. We argue that not controlling for the endogeneity of this variable leads to biased estimation of the impact of breastfeeding on the risk of childhood obesity[8].

Senauer and Kassouf (2000) represent one of the few studies that controls for endogeneity in the breastfeeding decision. The purpose of their study is to determine if breastfeeding leads to a better health outcome for children and if it further reduces the predicted demand for medical care. Estimating the impact of breastfeeding using binomial probit analysis, the authors found compelling evidence of lasting health benefits of breastfeeding on children's health. According to their results, breastfeeding reduces the probability of being sick by 15%. They also found benefits of breastfeeding to last more than six months. Following the same argument, we study the impact of breastfeeding on the risk of childhood obesity.

A study by Albino Barrera (1991) also argues that treating breastfeeding in a linear specification could lead to misleading results. Not accounting for a child's health endowment and family choices about goods and activities that impact child health could bias the relationship between breastfeeding and child health. Barrera used child's height for age score as the

---

[8] This study does not differentiate between overweight and obese. BMI above 95$^{th}$ percentile of child's age and sex is defined as being obese.

dependent variable and controlled for mother's height as a proxy for biological genes. There results show that, after accounting for maternal education, having a shorter duration of breastfeeding does not impact the child's health detrimentally. A paper by the Cebu Study Team (1992) also argues the fact that breastfeeding should be treated as an endogenous variable.

The objective of this research was to examine how breastfeeding affects obesity among children. This is a new approach to the issue, as previous literature has always treated the breastfeeding decision as being an exogenous variable in the context of childhood obesity. The breastfeeding decision highly depends on family preferences regarding the positive and negative effects of breastfeeding. It also depends on the availability of substitutes for breast milk and the quality of those substitutes that are available. Not controlling for family preferences and availability of substitutes for breast milk could lead to a biased estimated effect of breastfeeding on BMI. Given that this variable could be endogenous, Ordinary Least Squares (OLS) is inappropriate. To correct for this bias, this study used two-stage least squares and instrumental probit estimation technique to study the impact of breastfeeding on the BMI. This study used instruments to generate exogenous variations in the breastfeeding variable and estimate its impact on BMI[9] of the child.

Section I and II provide the necessary information about the sample used from NLSY, section III provides justification on the instruments used, and section IV presents the results followed by implications of the results.

---

[9] Compared to Rohrer index and weight for height index for children aged between 2-19 yr, BMI has been found to be the most efficient predictor of being overweight in children (Mei et al., 2002).

## II. Data Set

To investigate how breastfeeding relates to childhood obesity, I used year 1996 cross sectional data from the National Longitudinal Survey of Youth (NLSY) Child and Young Adult data set. The NLSY is a nationally representative survey which began in 1979 with 12,686 individuals. There were 6,283 women between ages 14 and 22 who participated in the initial survey year. The survey was conducted annually through 1994 and biennially thereafter. The NLSY Child and Young Adult data set presents year 1996 information on 8,120 children born to 6,283 women of NLSY 1979. Availability of pre-pregnancy and post-pregnancy information concerning mother and child makes this an ideal data set for this study.

This research project examined the effect of breastfeeding on two related dependent variables. The first dependent variable was body mass index (BMI), which was a continuous variable. BMI was calculated by dividing weight in kilograms by height in meters squared. This is the most commonly used measure of obesity. A BMI of 30 or above was categorized as obese for adults. For a child to be considered obese, he or she should rank above the $95^{th}$ percentile in terms of BMI among age and gender peers. The Center for Disease Control (CDC) growth charts from 2000 had been used to find $95^{th}$ percentile cutoff values for each age-gender combination among children in the data set. The second dependent variable was a discrete variable called "Obese." This variable equal's one if the child's BMI exceeds the 95th percentile among age and gender peers and zero otherwise. The distribution of the BMI variable is shown in Figure 2.1. It reveals an approximately normal, though slightly right-skewed, distribution of BMI values across the sample of children.

Various characteristics of the mother and child were treated as independent variables in this study. The NLSY also provides some information on other members of sampled families.

Controlling for maternal and pre-birth factors such as mother's age at time of childbirth, mother's pre-pregnancy weight, weeks of gestation, mother's education, smoking habits of mother, maternal employment after child's birth, family income, and whether the child's father is present in the household allows the model to explain much of the variation in the two dependent variables. This study also controlled for child characteristics such as gender, race, birth weight, whether the delivery was caesarian, and number of hours the child was left in childcare during his or her first year. The model used in this study is one of the most thorough models employed for studies on childhood obesity. Fortunately, the NLSY has sufficiently detailed information to estimate an accurate relationship between childhood obesity and breastfeeding.

## III. Instruments used

To correct for endogeneity, an instrumental variable approach was used. Instrumental variables were used to generate exogenous variations in the breastfeeding variable. Results from estimation techniques, probit and OLS, are presented to show the estimation differences when breastfeeding was treated endogenously rather than exogenously. The largest challenge in using an instrumental variable approach is to find instruments that satisfy the exogeneity and relevance conditions. A good instrument is highly correlated with the breastfeeding decision but not related to the error term of the model. Keeping this under consideration, and given the variables available in the data set, I used number of adults in each of four educational categories household as instruments. Four instruments have been used: number of adults with less than 12 years of education, number of adults with 12 to 13 years of education, number of adults with 14 to 15 years of education and number of adults with 16 or more years of education. Informed support from the father of the child, from other family members in the household, and from the social network (Raj et al., 1998; Matthews et al. 1998) act as encouragement for the mother to breastfeed the child. The benefits of breastfeeding on health outcomes of the child are widely established. Having more educated people in the household who understand the benefits of breastfeeding to the child makes the mother more likely to breastfeed. First stage results from two-stage least squares and instrumental probit estimation supported the argument that having more adults in the household with 14 or more years of education significantly increased the chances that the mother breastfeeds her child. On the other hand, having more educated people in the household reduces the probability that family members stay home to take care of the child. Their opportunity cost of staying home with the child increases with increased education. Therefore, an increased education level of other adults in the household decreases the extent to which these individuals impact the daily eating and physical activity decisions of the child. In most cases, it is the mother who controls such small, repetitive decisions. She may get advice and encouragement from other family members, but, in the end, she makes most decisions regarding such activities. This is true until the child is five or six years old. Once the child starts to choose things for himself-or herself, maternal control of his or her activities begins to diminish. The data set used for this study has children ranging from newborn to 15 years old. Any child who is above 10 years old is assumed to behaves according to his own preferences, which reduces the impact of maternal choices.

The mother's own preference regarding breastfeeding plays a major role in the decision of whether to breastfeed the child. Mothers who are conscious of their body or want to stay slim are less likely to breastfeed and thus more likely to rely on the available substitutes for breast milk (Wosje et al., 2004). Preferences seem to be independent of education or employment status and are likely related to the mother's own childhood circumstances or the culture she grew up in. If the mother decides against breastfeeding, she relies upon the available substitutes for breast milk. The quality of the substitutes available is something that is difficult to control. Poor quality substitutes could worsen the health of the child (i.e., increase BMI).

The F-statistic from the first stage, the Anderson canon, the correlation test, and the Hansen J test were each used to check the relevance and exogeneity of the instruments of this analysis. The first stage F-statistic value for the joint significance of instruments, 17.73 {F (4, 1246) = 17.73 p-value = 0.0000}, serves as an approximation as to the IV estimate quality. Given that the F-statistic is greater than 10, we can reject the hypothesis that first-stage instruments are jointly equal to zero (Staiger and Stock, 1997). According to the Hausman test, one cannot reject the hypothesis that breastfeeding is endogenous (p-value = 0.000). Therefore, a 2SLS method is preferable to OLS in that it produces estimates that are more consistent.

## IV. Estimation results

The specific models used for this analysis are:

$$BMI = \alpha_B + X\lambda_{\mathrm{B}} + Z\gamma_{\mathrm{B}} + \beta_B BF + \varepsilon_B$$

$$Obese = \alpha_o + X\lambda_o + Z\gamma_o + \beta_o BF + \varepsilon_o$$

Body Mass Index (BMI), a continuous variable, is a non-linear ratio of a person's weight and height; Obese is a dummy variable that uses a child's BMI to determine whether the Center for Disease Control considers him or her obese; BF measures whether the child was breastfed; and $\varepsilon$ is the standard error term. Table 2.1 presents a description of all the variables used in the study. Table 2.2 presents mean, standard deviation, minimum, and maximum values for each variable used. Table 2.3 presents the estimation results derived from OLS and instrumental variable results from two-stage least squares estimation. Breastfeeding reduces BMI in children by 0.496 points compared to children who were not breastfed. Higher family income was positively related to {0.013 (OLS) and 0.025 (IV)} an increase in BMI. Child's age was negatively correlated to BMI. OLS estimation showed that as children grow older in age, their BMI is significantly reduced by 0.024 per year. After controlling for endogeneity of breastfeeding, this impact slightly increased to 0.028. The effect of child's age on BMI could be picking up on some body growth factors. According to OLS estimation, being first born increased BMI by 0.236. Lack of experience of parents in taking care of the child could be a possible explanation for this increase. In instrumental variable estimation, the effect of being first born on BMI (1.067) is even greater than in OLS estimation.

Increase in maternal pre-pregnancy weight (in pounds) increases the child's BMI by 0.034 in the OLS regression and by 0.028 in the instrument variable regression. Maternal weight captures the 'thrifty gene' impact on the child's BMI.[10] Mothers who are overweight are more likely to have overweight children. Method of delivery appears to significantly affect the BMI of the child. Children who were born by caesarian section on average had a BMI that is 0.901 lower in OLS estimation and 1.356 lower in instrumental estimation, *ceteris paribus*. The number of childcare arrangements the child attends in the first year significantly affects the BMI of the child. This study found that the number of childcare arrangements was positively related to the BMI of the child.

[10] The thrifty gene hypothesis claims that parents of a thrifty genotype (i.e., ones who store fat more efficiently) are more likely to pass the gene to offspring (Connor, 2003).

Various other explanatory variables were used in the analysis. Given the p-values, this study finds no significant effect of the remaining variables on BMI. Previous literature has found maternal employment to have a positive significant impact on a child's BMI (Anderson et al., 2003; Ruhm, 2004; Finkelstein et al., 2005). However, the results from the sample used for this study did not show any significant effect of maternal employment on the child's BMI. Though insignificant, the maternal employment coefficient is negatively correlated with BMI. This negative correlation could be due to some unobserved maternal characteristics (Anderson et al., 2003). Anderson et al. 2003 also discuss the possibility of this variable being endogenous. Given the focus of this study, I did not treat this variable as endogenous. Maternal education is another variable which, though insignificant, is sometimes thought to affect child BMI. Maternal education is negatively related to the BMI of the child. Educated mothers are more aware and perhaps have a better understanding of the things related to a better child health outcome. Neither of these variables showed any significant effect upon BMI of the child.

This study also analyzed the relationship between endogenous breastfeeding and the probability of the child being obese using instrumental probit analysis. Table 2.4 presents the results from the probit estimation with the dummy variable ''Obese'' as the dependent variable. Marginal effects from the probit model are also shown in Table 2.4 since they are easier to interpret in such a model. Table 2.5 presents the results after treating breastfeeding as an endogenous variable. Being breastfed reduces the probability of the child being obese by 17.9%. Though we saw a similar effect of breastfeeding on reducing the probability of childhood obesity as of breastfeeding on BMI in 2SLS, the impact was not significant in the instrumental probit analysis. If the mother was overweight, the child is 0.2% more likely to be obese. The mother's weight significantly affects child BMI, as well as the probability that the child will be obese.

The results from probit analysis also showed that as the child grows older, the probability of being obese declines by 0.1%.

## V. Conclusion

Instrumenting for breastfeeding, to control for endogeneity in the breastfeeding variable reduces BMI by 5.669. While this may appear to be a large marginal decline in BMI, previous literature shows a similar drop in BMI when the effect of breastfeeding is correctly estimated. Von Kries (1999) showed that breastfeeding reduces the risk of obesity by 57%, and Liese et al. (2001) found that breastfeeding could reduce the risk of being overweight by 71%. This study found a positive correlation between family income and BMI. With the increasing popularity of such normal goods as cable television and video games, the lifestyle of the high income family child may have become more sedentary. Such leisure activities for high income children have been found to cause a higher BMI in children, *ceteris paribus* (Vandewater, Shim, and Caplovitz (2004). The child's age was negatively correlated with BMI. The National Health and Nutrition Examination Survey (NHANES) from 2003-2004 reported 18.8% obesity among children aged 6-11 years and a 17.4% level of obesity among children aged 12-19 years. This may be partly attributable to large height gains in adolescent years. Also, older children are more likely to be involved in rigorous sports training and have more independence compared to younger children. Younger children must rely mostly on family members to take them out to play whereas older children can engage in activity on their own. A negative impact of being firstborn on BMI could be explained by lack of experience on the parental side.

Strauss and Knight (1999), Locard et al. (1992), Whitaker et al. (1997), and many others have shown how parental obesity is a significant factor in determining the BMI of the child. All these studies show at least a two-fold increase in the risk of childhood obesity if parental obesity is present in the family. This study controlled for the mother's weight to capture its impact on the child's BMI and also on whether the child is obese or not. Maternal weight also captures the eating and home environment of the family. Families that have healthier lifestyles are less likely to have obese children. Healthier lifestyles could include things such as exercise routines and nutritious eating habits. Hediger et al. (2001) showed that maternal weight is the strongest predictor of the child's weight. Using a cross sectional data set from NHANES III survey and controlling for factors such as race, education of the family members, mother's age at child's birth, birth order, maternal smoking habits, gender of the child, birth weight, length of gestation, mother's weight before pregnancy, and duration of breastfeeding, the authors were unable to find any dose-response relationship between breastfeeding and BMI. However, they found maternal weight to be the most significant predictor of a child's BMI. In the current study, after controlling for very similar confounding variables, our results showed that maternal weight and breastfeeding are both strong predictors of a child's BMI. OLS regression results are similar to what Hediger et al. (2001) found. Maternal weight is a significant predictor of BMI, and breastfeeding is an insignificant predictor of a child's BMI. However, after controlling for the endogeneity of the breastfeeding variable, this study found breastfeeding and maternal weight to be very strong predictors of child's BMI.

Medical literature shows mixed results regarding the relation between BMI and method of delivery (Silva et al., 1998, Malloy et al., 1989). Some prior studies show that, after controlling for confounding factors such as maternal education, family income, length of gestation period, and cost of the medical procedure, method of delivery no longer significantly affects the BMI of the child. Explanatory variables used in this study control for a large variety of family background variables. However, the model still reveals a significant impact of method of birth on BMI. A child going to multiple childcare arrangements could be indicative of various family and maternal characteristics. Blau and Rabins (1991) showed that turnover in childcare arrangement was positively related to maternal characteristics. Few arrangements in the first year indicated that the mother is more likely to have planned for the child to have consistent care. Such a mother is expected to be more attentive to the child's well-being. Previous research also shows a positive association between hours spent at childcare and BMI (Story et al., 2006). The positive association found in this study could be due to the kind of food provided and the amount of physical activity the child has each day in childcare.

Results from this study provide helpful insights to policymakers and concerned parents about the positive benefits of breastfeeding on the weight of the child. Optimal maternal leave policies depend upon an understanding as to the true benefits of breastfeeding. In this sense, the relationship between breastfeeding and obesity has important microeconomic policy implications. After correcting for any estimation bias due to the breastfeeding variable being endogenous, this study shows the benefits of breastfeeding and the rationale behind promoting such a decision to concerned parents. Given these results, longer maternity leaves or more comfortable work situations might be advisable to lower the costs of breastfeeding to the mother.

# VI. Figures and Tables

## Table 2.1 Data description, 1996 National Longitudinal Survey of Youth (NLSY)

| Variables | Description of the variable |
|---|---|
| *Dependent Variables* | |
| BMI of child | Body Mass Index of the child |
| Obese | 0/1: if the child is obese |
| | |
| *Independent variables* | |
| Child's age | Age of the child in months |
| First born | 0/1: if the child is first born |
| Income | Total net family net income |
| Mothers age at birth | Age of mother at child's birth in years |
| Mother grade | 0/1: Mother is high school graduate |
| Hours worked | Number of hours worked per week by mother in the first quarter after child's birth |
| Smoking | 0/1: Mother smoked 12 months before birth |
| Mother's Weight | Weight of mother before pregnancy in pounds |
| Cesar | 0/1: Was child delivered by cesarean section |
| Birth weight | Birth weight of child in ounces |
| Father present | 0/1: Father present in the household |
| Gestation | Length of gestation of child in weeks |
| Breastfeeding | 0/1: if the child was ever breastfed |
| Child care | Number of child care arrangements for child in 1st year |
| adult 11 | number of adults in the household with highest grade completed less than 12 |
| adult 12-13 | number of adults in the household with highest grade completed = 12-13 years |
| adult 14-15 | number of adults in the household with highest grade completes= 14-15 years |
| Male | 0/1: if the child is male |
| Hispanic | 0/1: if the child is hispanic |
| Black | 0/1: if the child is back |
| Non-race | 0/1: if the child is non-black and non-hispanic |

**Table 2.2 Summary Statistics, 1996 National Longitudinal Survey of Youth (NLSY)**

| Household, mother and child characteristics | Mean (S.D) | Min. | Max. |
|---|---|---|---|
| BMI | 23.391(7.443) | 1.82 | 77.46 |
| Breastfeeding | 0.442(0.497) | 0.00 | 1.00 |
| Child's age | 125.721(65.008) | 0.00 | 304.00 |
| First born | 0.351(0.477) | 0.00 | 1.00 |
| Income | 47787.92(53007.28) | 0.00 | 150000.00 |
| Mothers age at birth | 24.504(5.230) | 10.00 | 40.00 |
| Mother grade | 0.803(0.397) | 0.00 | 1.00 |
| Hours worked | 34.886(11.741) | 0.00 | 96.00 |
| Smoking | 0.332(0.471) | 0.00 | 1.00 |
| Mother's Weight | 133.465(28.07) | 67.00 | 520.00 |
| Cesar | 0.210(0.408) | 0.00 | 1.00 |
| Birth weight | 116.003(21.708) | 0.00 | 268.00 |
| Father present | 0.575(0.494) | 0.00 | 1.00 |
| Gestation | 38.623(2.285) | 18.00 | 51.00 |
| Child care | 1.256(0.631) | 0.00 | 10.00 |
| adult 11 | 0.477(0.757) | 0.00 | 6.00 |
| adult 12-13 | 0.935(0.853) | 0.00 | 6.00 |
| adult 14-15 | 0.254(0.496) | 0.00 | 3.00 |
| adult 16 | 0.302(0.625) | 0.00 | 5.00 |
| Male | 0.510(0.499) | 0.00 | 1.00 |
| Hispanic | 0.189(0.392) | 0.00 | 1.00 |
| Black | 0.279(0.449) | 0.00 | 1.00 |

**Table 2.3 National Longitudinal Survey of Youth 1996: OLS and 2SLS Estimates**

| | OLS | | IV 2SLS | |
|---|---|---|---|---|
| *Dependent Variable: Body Mass Index(BMI)* | Coefficient | t-statistics | Coefficient | t-statistics |
| Breastfeeding | -0.496 | 1.155 | -5.669*** | 2.670 |
| Child's age | -0.025** | 2.574 | -0.028*** | 2.723 |
| First born | 0.236 | 0.579 | 1.067** | 2.008 |
| Income | 0.001 | 1.243 | 0.001* | 1.818 |
| Mothers age at birth | -0.168* | 1.684 | -0.113 | 1.068 |
| Mother grade | -0.957 | 1.092 | -0.671 | 0.768 |
| Hours worked | -0.006 | 0.332 | -0.021 | 1.024 |
| Smoking | 0.377 | 0.757 | -0.096 | 0.181 |
| Mother's Weight | 0.034*** | 4.871 | 0.028*** | 3.506 |
| Cesar | -0.901** | 2.134 | -1.356*** | 2.709 |
| Birth weight | 0.001 | 0.107 | 0.004 | 0.290 |
| Father present | 0.251 | 0.577 | 0.397 | 0.844 |
| Gestation | -0.023 | 0.211 | 0.069 | 0.554 |
| Child care | 1.174*** | 2.772 | 1.301*** | 2.912 |
| Male | -0.527 | 1.276 | -0.437 | 1.011 |
| Hispanic | 0.320 | 0.612 | 0.057 | 0.101 |
| Black | 0.775 | 1.328 | -0.356 | 0.517 |
| Constant | 25.413*** | 5.085 | 23.688*** | 4.522 |
| N | 1267 | | 1267 | |
| Hausman chi-square | | | 44.43 | |
| P-Value | | | 0.000 | |
| F-test for Joint Significance of Instruments | | | 17.73 | |
| P-Value | | | 0.000 | |
| Hansen over-identification test: J statistic | | | 0.688 | |
| P-Value | | | 0.8761 | |

**Table 2.4 National Longitudinal Survey of Youth 1996: Probit Estimates**

| | Probit Estimates | | |
|---|---|---|---|
| *Dependent Variable: Obese* | Marginal effect | Coefficient | z-statistics |
| Breastfeeding | -0.002 | -0.005 | 0.068 |
| Child's age | -0.001 | -0.003* | 1.915 |
| First born | -0.004 | -0.011 | 0.138 |
| Income | 0.000 | 0.000 | 0.264 |
| Mother's age at birth | -0.006 | -0.016 | 0.912 |
| Mother's grade | -0.006 | -0.017 | 0.121 |
| Hours worked | 0.003 | 0.007* | 1.826 |
| Smoking | 0.008 | 0.022 | 0.253 |
| Mother's weight | 0.002 | 0.006*** | 4.519 |
| Cesar | 0.046 | 0.117 | 1.439 |
| Birth weight | -0.000 | -0.001 | 0.629 |
| Father present | 0.013 | 0.035 | 0.405 |
| Gestation | -0.004 | -0.011 | 0.522 |
| Child care | -0.003 | -0.009 | 0.148 |
| Male | -0.189 | -0.047 | 0.659 |
| Hispanic | 0.011 | 0.028 | 0.281 |
| Black | 0.004 | -0.010 | 0.106 |
| Constant | | 0.222 | 0.240 |
| N | 1267 | | |

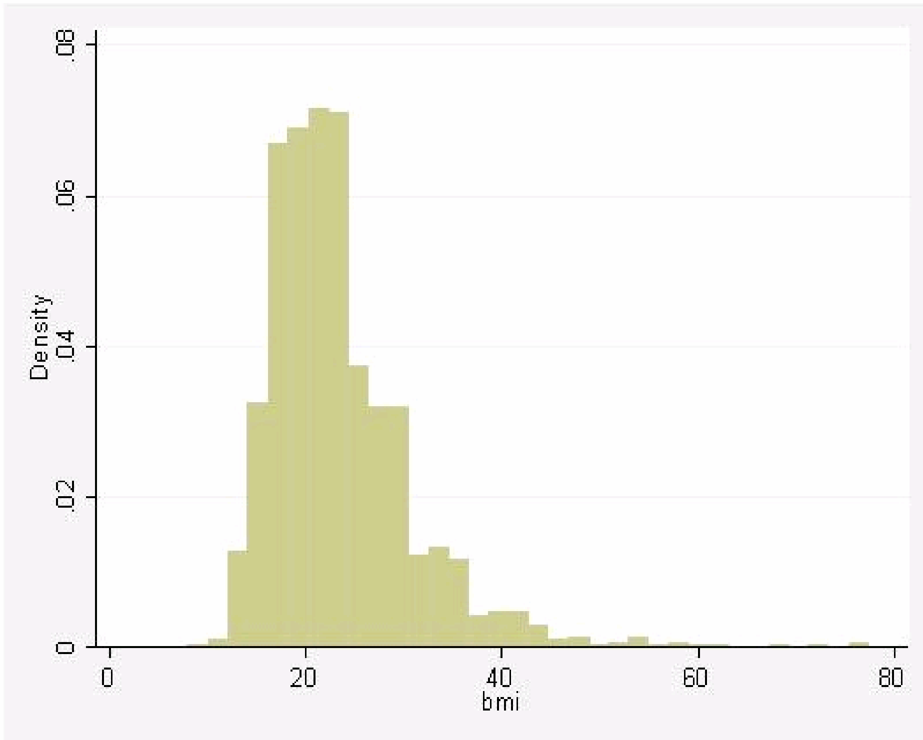
Note: * indicates p<0.10, ** indicates p<0.05, *** indicates p<0.01

**Table 2.5 National Longitudinal Survey of Youth 1996: Instrumental Probit Estimates**

| | IV Probit Estimates | | |
|---|---|---|---|
| *Dependent Variable: Obese* | Marginal effect | Coefficient | z-statistics |
| Breastfeeding | -0.179 | -0.455 | 1.310 |
| Child's age | -0.001 | -0.003** | 2.021 |
| First born | 0.025 | 0.062 | 0.655 |
| Income | 0.000 | 0.000 | 0.588 |
| Mother's age at birth | -0.004 | -0.011 | 0.621 |
| Mother's grade | 0.003 | 0.008 | 0.057 |
| Hours worked | 0.002 | 0.006 | 1.418 |
| Smoking | -0.007 | -0.019 | 0.202 |
| Mother's weight | 0.002 | 0.005*** | 3.907 |
| Cesar | 0.030 | 0.077 | 0,884 |
| Birth weight | -0.000 | -0.001 | 0.511 |
| Father present | 0.019 | 0.048 | 0.543 |
| Gestation | 0.001 | -0.003 | 0.131 |
| Child care | 0.000 | 0.002 | 0.038 |
| Male | -0.016 | -0.040 | 0.543 |
| Hispanic | 0.002 | 0.005 | 0.053 |
| Black | -0.043 | -0.108 | 0.888 |
| Constant | | 0.073 | 0.077 |
| N | 1267 | | |

Note: * indicates $p<0.10$, ** indicates $p<0.05$, *** indicates $p<0.01$

**Figure 2.1 Distribution of Body Mass Index (BMI) Variable**




Data Source: National Longitudinal Survey of Youth 1996

# CHAPTER 3 - Differences in Student Evaluations of Principles and Non-Principles Economics Courses and the Allocation of Faculty across these Courses

## I. Introduction

As a discipline, economics focuses on the allocation of resources. In universities, one of the most important allocation decisions is matching faculty with the courses the department offers. The public may prefer that the best teachers be assigned to principles courses, so that these faculty have contact with the most students, but faculty may prefer upper-level courses because of their smaller class size and greater interest on the part of students. Another possible reason to avoid introductory classes is if student evaluations of teaching (SET) are lower in principles of economics than in upper-level economics courses and if these differences are not adequately accounted for by the administrators who evaluate faculty for salary increments, tenure, promotion, and teaching awards.

This study provides perspective on student assessment of principles and non-principles courses and on the allocation of faculty across these courses. It does so by examining teaching evaluations over 19 semesters in the department of economics at a large public university. It compares evaluations of principles and upper-level undergraduate courses in economics and asks whether differences in raw SET scores can be explained by differences in class size, grading policy, and other characteristics. It also considers whether the relationship between SET ratings and the explanatory variables is the same for principles and non-principles courses.

The models are estimated first by ordinary least squares, the technique most commonly used in the literature, and then after allowing for faculty fixed effects. One reason to allow for fixed effects is that faculty may not be randomly assigned to principles and non-principles classes. If the best (worst) teachers are assigned to principles classes, then OLS estimates understate (overstate) the amount by which a given faculty member can expect to see her SET score fall if she switches from a non-principles class to principles. The problem is that ordinary least squares cannot disentangle the separate effects of teacher and course. More generally, personality and other unmeasured faculty characteristics may be correlated with the model's variables, which can bias estimates, obtained by ordinary least squares.

Another advantage of the fixed-effects approach is that it provides an alternative to average SET scores for assessing faculty teaching. Based on this metric, we study the distribution of teacher effectiveness in the department. We examine the correlation between individual faculty effects in principles and non-principles courses to provide an assessment of whether highly rated non-principles teachers also tend to be highly rated principles teachers. If some teachers have a comparative advantage in principles courses and others in upper-level classes, the correlation need not be high. We also address whether faculty who teach principles are concentrated among the highest rated or the lowest-rated teachers in the department.

The essay proceeds along the following lines. We start with a brief review of studies that estimate SET equations separately for principles and non-principles courses. Next, we describe economics classes at Kansas State University and analyze teaching data at this university. The empirical model is formulated, alternative specifications are estimated and interpreted, and results are compared with the literature. After a section on faculty effects and the distribution of faculty across courses, we summarize the essay's findings.

## II. Do Students Evaluate Principles and Non-Principles courses Differently?

Typically, studies that assess economics instruction do so for principles or collectively for all economics courses; but there are exceptions. In the study closest to ours, McPherson (2006) uses a fixed-effects model to study teaching evaluations over 17 semesters at the University of North Texas (UNT). Estimating equations separately for principles and upper-division classes, he finds that greater teaching experience is associated with better SET ratings in principles but not in non-principles courses and that a larger class size has a significant adverse effect only in principles classes. His raw data indicate that SET scores are marginally higher in upper-division classes at UNT (by .12 on a scale of 1 to 4).

Boex (2000) uses ordered probit equations to study student evaluations at Georgia State University. He presents results separately for core and non-core courses. At the undergraduate level (the focus of the present study), the core category contains both principles and non-principles courses, but Boex reports that most of the core observations come from principles classes. Empirical estimates suggest that the structure of equations may differ for the two categories of courses. For example, class size and response rate are significant determinants of teaching evaluations only in the core classes, and student motivation is quantitatively more important in non-principles classes. Boex finds that the raw SET score is lower in core classes (3.86 versus 4.08), but he does not test whether the difference remains after controlling for other factors.

Weinberg, Fleisher, and Hashimoto (2007) study teaching evaluations for principles of microeconomics, principles of macroeconomics, and intermediate microeconomics at Ohio State University. Their data indicate differences in the characteristics of principles and non-principles teachers. For example, mean teaching experience at the university was 4.8 years for intermediate microeconomics compared to 16.0 and 16.4 years, respectively, for principles of microeconomics and principles of macroeconomics. Graduate students were assigned to 23 percent of the intermediate classes and to 12 percent and 16 percent of the principles classes. Foreign-born instructors taught 33 percent of the intermediate classes compared to 16 percent and 21 percent of the principles classes. The authors did not pool data to test whether there was a common structure to overall SET ratings across the three courses, but differences in estimated coefficients are small relative to reported standard errors, so a common structure cannot be ruled out.

Among studies that estimate a single SET equation for all economics courses, a dummy variable is sometimes added to allow for differences across categories of courses. Typically, the coefficient for principles or a principles-related variable is negative but not statistically significant. This is the pattern for introductory courses (Nelson and Lynch 1984), core courses (Krautmann and Sander 1999; Isely and Singh 2005), and lower-division courses (Nichols and Soper 1972). Similarly, Aigner and Thumb (1986) find a negative and insignificant coefficient for introductory classes. But of greater interest, the coefficient of the interaction of this variable with class size is significantly less than zero. This finding suggests that the effect of class size, and perhaps other variables, may differ for principles and non-principles courses.

## III. Economics Classes at Kansas State University

***Background***

Kansas State University is a public university with enrollment of over 20,000 students. Although the university offers master's and doctoral degrees in economics, this study focuses exclusively on undergraduate instruction in economics. The department offers separate courses in principles of macroeconomics and principles of microeconomics, both at the 100-level. All other economics courses are, by their course number, considered upper-level classes. Most colleges on campus require a principles course of their majors, commonly principles of macroeconomics. Less than one percent of the students who take principles are economics majors when they sign up for the course.

In upper-level courses, the primary sources of students are the College of Business Administration, which requires two economics courses beyond principles, and the economics major (which is in the College of Arts and Sciences). Even though economics and business students must complete a certain number of economics courses, students generally have discretion over which courses they take, except that intermediate theory and senior seminar are required of all economics majors. Therefore, the percentage of students taking a course because they find the subject interesting is likely to be greater for non-principles courses.

The two types of courses also differ in other dimensions. To accommodate the colleges that require principles, these classes are taught in lecture halls. There are no discussion sections, so the classes emphasize the lecture format. Exams tend to be all or predominantly multiple-choice in nature, and most homework assignments are graded online. Term papers are not assigned, and given the large class size, it is rare for any written assignment to be made. As such, the principles classes can be characterized as “chalk and talk” (Becker and Watts 1996). In contrast, non-principles classes are much smaller, typically involve greater student interaction, require term papers or similar assignments, and rely heavily on an exam format other than multiple-choice.

Faculty teaching assignments are negotiated. Because faculty are not forced to teach principles, the department relies on graduate students to cover principles classes not claimed by faculty. When the department chair indicated at a faculty meeting that he would like to see additional faculty teach principles, one of the responses was that teaching principles would likely not be in the best interest of individual faculty. Some faculty voiced their opinion that teaching evaluations tended to be lower in principles, perhaps because of lower student interest or larger class size, and that the department did not adequately take this point into account when evaluating faculty teaching. One goal of the present essay is to learn the extent to which SET scores vary by type of economics class and to obtain a sense as to what would be an equitable adjustment for faculty who teach principles—if in fact, any adjustment is appropriate.

Toward that end, the authors solicited full-time faculty for permission to access their teaching records. Even though records are housed in the departmental office, teaching evaluations belong to individual faculty, so their approval was necessary to proceed. Faculty were promised that no attempt would be made to identify individual teachers. To further assure faculty that data would remain anonymous, various procedures were put in to place to protect the identity of faculty. Although the bulk of faculty teach two courses per semester, faculty with less than a 40 percent research weight teach three courses. So that these faculties could not be identified, teaching evaluations were obtained for at most two courses per semester.[11] As an added privacy screen, each faculty member was matched with at least one other person, based on length of time at the university, so that this variable could not be used to identify individuals.[12]

[11] Data are available for less than two courses when the teaching assignment included graduate courses, when the faculty member had an administrative appointment or bought out of teaching with external funding, and when the faculty member was on leave.
[12] As an example, if one faculty member left the university after three years in the sample and a second left after four years, data for the fourth year were not collected for the second faculty member.

The sample period consists of 19 semesters, spring 1997 through spring 2006. During this time frame, 26 different full-time faculty taught for the department. All 26 faculty were solicited by e-mail for permission to include their data in the study. They were directed to indicate to an administrative assistant in the department whether or not they would allow us access to the data. After one month, the administrative assistant contacted anyone who had not responded. After this second contact, 24 of the 26 faculty members granted permission to use their data.

## III. The Data

The data set consists of 284 classes taught by these 24 faculty. Most of the data came directly from the standard student evaluation form used at Kansas State University (TEVAL), which is administered by the Center for the Advancement of Teaching and Learning, but other sources supplemented these data. The TEVAL form does not provide information about the student's current GPA or grade expected in the course. (The only TEVAL statement pertaining to grades is the following: "The instructor's grading procedures were fair and equitable.") Therefore, we asked the office staff to add data on actual class GPA for each of the classes in the sample (which the department already had compiled). In addition, the staff constructed the teaching experience variables.

In order to see if teaching evaluations improved as the instructor gained experience, we provided departmental staff with information on the faculty member's start date at the university for each of the 26 faculty members potentially in the sample, and the staff created four variables for teaching experience for the 24 faculty actually in the sample. *Year 1* designates the faculty member's first year at the university, meaning that the person is in her first or second semester of teaching. *Year 2* signifies that the person is in her second year at the university, and *Year 3-4* indicates third or fourth year of teaching. The reference category, *Year 5+*, designates that the person has previously completed at least four years of teaching at the university.

Table 1 presents summary statistics for the 284 classes in the sample, separately for principles and non-principles courses. The table contains separate variables for actual GPA of the class, which is used only in supplemental regressions, and *Expected GPA*, a variable that is defined later in the essay. The dependent variable, *Teacher Effectiveness*, is defined as "overall effectiveness as a teacher." It is considered the primary measure of teaching quality in the TEVAL survey.

The table reveals that average effectiveness, as judged by students, is higher in non-principles classes than in principles classes, 3.91 versus 3.61 on a scale of 1 (low) to 5 (high). Teaching experience is greater in the non-principles classes: 81 percent of these classes were taught by a teacher in at least her fifth year at KSU, compared to 48 percent for principles classes; 17 percent of the principles classes in the sample were taught by a first-year faculty member, compared to only 5 percent of non-principles classes. Student characteristics also differ. *Student Interest*, defined as "interest in the course before enrolling," is much higher in non-principles (3.35 versus 2.89), and *Student Effort* ("effort to learn in the course") is slightly higher (3.75 versus 3.64).

Principles classes are graded more harshly. Their GPA is 2.30 on a 4.0 scale, whereas the GPA in non-principles classes, while low by university standards, is 2.74. On the other hand, principles classes are not viewed as graded less fairly. The average for the question on grade equity is 4.0 for both types of courses. *Grades Equitably* is capturing a dimension of grading other than easy grading. In fact, for both principles classes and non-principles classes, *Student*

*Effort* is correlated more highly with *Grades Equitably* than is *GPA*. Students indicated to the authors that they interpret the equity question in terms of fairness: relative grades corresponding to mastery of the material. In some cases, this may mean giving partial credit for exam answers; in general, it means a close correspondence between grades of individual students and the grades they deserve.[13] Whereas class GPA measures the central tendency of an instructor's grades, equity refers more closely to the distribution of grades across students.

Another key difference between the types of courses is number of students enrolled. Class size averages 148 for principles classes compared to 41 for non-principles classes. The response rate is also much lower among principles students, .56 versus .77. Reasons may include the greater anonymity in large classes (the teacher is unlikely to know that you skipped class the day of the evaluation) or lower interest in the subject matter. Differences in response rates also raise the possibility of selectivity biases (Becker and Watts 1999; Becker and Powers 2001). The sample of students who fill out the questionnaire is a censored sample of the population of students who enrolled in the class. Therefore, students who attend class when evaluations are administered and take time to fill them out may differ, in their assessment of teaching, from students who do not fill out the questionnaires. In that event, the average value of students' assessments depends on the fraction of students who respond.[14]

---

[13] At the request of the instructor, one of the authors presented this paper to students in Senior Seminar, a class limited to economics majors. Prior to the presentation, he asked students to indicate how they interpreted the equity question. One student indicated that equity means that the teacher follows carefully the terms of the course syllabus and does not make exceptions, e.g., extra credit for certain students. A second person offered the view that equitable grading means grades that are proportionate to the amount learned. If a student misses the numerical answer on a question but demonstrates that he understands the underlying concept, he should receive appropriate partial credit. When asked if they associated equitable grading with easy grading, all 21 students present said "no." When asked if it meant "receiving the grade deserved," an expression used by one of the students, 18 of the students responded in the affirmative. In a second class, students added that another component of equity is testing based on the material taught in class.

[14] One could potentially adjust for selectivity using the approach of Heckman (1979), as did Becker and Powers in their study of student learning; but that requires data on non-respondents, which we do not have.

## IV. Specifying the Model

Our model builds on an extensive literature on student assessment of teaching. In particular, we estimate the following equation:

$$Teacher\ Effectiveness = X_{cit}\beta + \mu_i + \varepsilon_{cit} \quad (1)$$

where $X$ is a vector of variables that influence student assessments of teaching. The subscript $c$ refers to the particular class taught by professor $i$ in semester $t$ that is being evaluated. Because up to two classes are evaluated each semester for faculty in our sample, $c$ = 1 or 2. The error term consists of two components: a faculty fixed effect ($\mu_i$) and the classical error term ($\varepsilon_{cit}$), representing white noise. For purposes of comparison, the model is also estimated by ordinary least squares. Equation (1) is estimated both for the pooled sample, which includes principles and non-principles courses, and separately for each category.[15]

An advantage of allowing for faculty fixed effects is that they control for unmeasured faculty characteristics that may be correlated with the variables in the $X$ vector.[16] In that event, OLS estimates will be biased. Fixed effects also help control for the nonrandom assignment of faculty to principles and non-principles courses. If principles teachers tend to be either more effective or less effective than non-principles teachers, then the *Principles* variable in the pooled regression will confound the effect of teacher quality and type of course when the model is estimated by ordinary least squares. By controlling for individual teacher, by seeing how student evaluations differ between principles and non-principles classes for the same teacher, the fixed-

[15] Potentially, equation (1) could also be estimated for more disaggregated groups of courses; but, to assure faculty that they could not be identified, we restricted the analysis to principles and non-principles courses, which we and other members of the department judged to be the key distinction.

[16] Apart from personality, unmeasured characteristics include such things as command of English, which may be important in light of studies that find lower student ratings for teachers who are not proficient in the English language (Finnegan and Siegfried 2000, Bosshardt and Watts 2001, and Saunders 2001). English skills might reasonably be correlated with student effort, response rate (and attendance), and class size.

effects model provides a clearer picture of the relationship between course category and student rating.

Baseline variables in the *X* vector include the years-of-teaching variables previously defined (*Year 1*, *Year 2*, and *Year 3-4*); information on student characteristics (*Student Interest* and *Student Effort*); a quadratic specification of class size; *Response Rate*; expected class GPA (discussed below); dummy variables for each of 18 semesters to allow for period effects; and, in the version that pools principles and non-principles classes, a variable that indicates that the class is principles of economics (*Principles*). In a second version of the model, we add *Grades Equitably*, a variable missing from prior studies.

Previous research suggests that student assessments depend positively on student interest in the class and on work effort (see Marsh and Duncan 1992 and Boex 2000). Commonly, the estimated effect of teaching experience is insignificant (as in Feldman 1983 and Weinberg, Fleisher, and Hashimoto 2007), but this finding may be sensitive to specification of teaching experience. Using a continuous variable may miss any effect that is limited to rapid early gains, as new teachers adapt to the students and learn how to pull up ratings. Our specification of teaching experience allows us to test for this possibility.

The effect of class size in the literature is mixed. Among the studies that find no effect are Nichols and Soper (1972), Nelson and Lynch (1984), Krautmann and Sander (1999), and Finegan and Siegfried (2000). Some studies find a *positive* relationship between class size and SET ratings (Mirus 1973; Boex 2000—for core courses only). Other studies find the expected inverse relationship (Isely and Singh 2005; Bedard and Kuhn 2005; and McPherson 2006). As Bedard and Kuhn point out, a potential limitation of class size variables in cross-sectional studies is that the effects of class size and instructor quality may be confounded if the best teachers tend to be assigned to either large or small classes. That is not a problem, however, for fixed-effects estimates, which study variation in class size for a given instructor. In addition, Bedard and Kuhn find that the negative effect of greater class size tapers off, so that neither the linear specification nor the single class-size dummy that some studies are forced to rely on is appropriate. Accordingly, we adopt a quadratic specification.[17]

One of the most extensively studied relationships is that between course grades and teacher ratings. Although most studies find a positive relationship, there is no agreement as to what the relationship is capturing. Some authors argue that faculty are buying higher SET scores with more lenient grading, but others are not so sure. The positive relationship between grades and SET may reflect other influences. For example, grade may be correlated with an omitted variable, such as amount learned or teacher empathy, in which case the coefficient of the grade variable provides a biased estimate of the influence of teacher grading policy.

One approach is to estimate a model of two-stage least squares, but that requires obtaining identifying variables that determine student grade and that are uncorrelated with the error term of the SET equation. In many cases, the identifying variables of prior studies have been challenged (see Krautmann and Sander 1999). In any event, some studies that test for endogeneity find evidence of it (Seiver 1983; Nelson and Lynch 1984), whereas other studies conclude that expected grade can be treated as exogenous (Krautmann and Sander 1999; Isely and Singh 2005; and McPherson 2006).

Because of potential endogeneity of current grades and teacher effectiveness, our baseline model relies on a pre-determined measure of a teacher's grading standards—the expected GPA of the class before the semester begins, which is based on the faculty member's recent reputation

[17] A cubic specification did not improve the fit of the equation or alter results appreciably, so we prefer the more parsimonious quadratic formulation.

for grading in this type of course (principles or non-principles). For each type of course, *Expected GPA* is the GPA of a given faculty member the most recent time she taught the course. For the first semester of the sample, spring 1997, we use GPA values the year preceding our sample period (fall 1996 or, if these data are unavailable that semester, spring 1996). For a faculty member who joins the university during our sample period, there is no grading reputation to capture. For this first semester, *Expected GPA* is assigned the mean GPA of the department for this category of course in the prior semester. In other words, we assume that a student signing up for a course has no information on which to assess whether his instructor is an easy or hard grader, so he assumes that the GPA of her class corresponds to the departmental average for this type of course the last time it was taught. We also report (in a footnote) the coefficient of actual GPA in the present semester as a way to compare our estimates with those of studies that treat course grade as exogenous.

## V. Empirical Estimates

### ***Comparing OLS and Fixed-Effects Coefficients***

Initially, we estimated the baseline version of the model separately for principles and non-principles samples, but a Chow test indicated that we could not reject the null hypothesis of a common structure for both types of courses.[18] Therefore, columns 1-2 of table 2 provide results only for the pooled sample. Estimates are presented first for ordinary least squares

[18] When we allowed for different coefficients for *Year 1*, *Year 2*, *Year 3-4*, *Class Size*, *Class Size*$^2$, *Student Interest*, *Student Effort*, *Response Rate*, and *Expected GPA* (the model specification of table 2, column 2), we could not reject the hypothesis of identical coefficients for the two samples at even the .30 level of significance; $F(9, 223) = 1.10$. Even when we use ordinary least squares, which an $F$- test indicates is inappropriate, we are unable to reject the null hypothesis at the .10 level; $F(9, 246) = 1.44$.

(column 1) and then for the model with faculty fixed effects (column 2).[19] Even though an *F*-test indicates that the FE specification is preferred,[20] OLS estimates provide a basis for comparing our results with those of earlier studies and reveal the extent to which results are sensitive to the specification.

Regardless of the specification, the regressions indicate that SET ratings are lowest during the first year that a faculty member teaches for the department. After the first year, there is no evidence that ratings change with added experience. These results suggest that faculty are quick learners. They find out during their first year which approaches are effective and which are not, and they make the adjustments necessary to pull up their ratings. McPherson (2006) defined his low-experience variable more broadly, but his results are consistent with ours, at least for principles classes. At the University of North Texas, principles teachers were rated lower if they had less than five semesters of experience. For non-principles teachers, there was no evidence of an experience effect; but it is possible that an effect exists for low-experience teachers as defined in the present study, those with one or fewer semesters of prior teaching experience.[21]

Our results also indicate that student evaluations fall with class size, at least up to a certain point. Based on the estimates of column 2, student evaluations are minimized at a class size of 156, not far from the maximum class size of 175. According to the estimates of column 2, increasing class size from 20 to 40 students would reduce *Teacher Effectiveness* by .19, and

[19] We also estimated the models without semester effects, but in all cases an *F*-test of the hypothesis that semester dummies are jointly equal to zero could be rejected at the .01 significance level.

[20] The null hypothesis that individual faculty effects are jointly equal to zero is rejected at the .0001 level; $F(23, 231) = 9.60$. In addition, the Hausman test rejects the assumption of no correlation between $\mu_i$ and the explanatory variables at the .0001 level ($\chi^2 = 806.27$), an assumption required for random effects. Therefore, results support the fixed-effects specification over both ordinary least squares and random effects.

[21] Using aggregate, interdisciplinary data, Centra (1978) also found that SET ratings tend to be lowest during a teacher's first year.

doubling class size again, to 80, would reduce *Teacher Effectiveness* by an additional .29, other things equal. The predicted difference in *Teacher Effectiveness* for classes with enrollment of 41 students (the average for non-principles classes) and 148 students (the average for principles classes) is .49. This estimate is very similar to that predicted by the model of Bedard and Kuhn and underscores the importance of adjusting for class size.[22]

Somewhat surprisingly, teaching effectiveness is not significantly related to a student's interest in taking the class. Perhaps that is because students have a poorly defined sense of what an economics course will be like. Students who, before enrollment, are not especially interested in the course may find it more to their liking than they anticipated and rate the instructor as highly as students whose interest level was high.[23] Similarly, students who (before the start of the class) thought they were interested in the course may learn during the semester that they are less excited about the material than they expected and, therefore, rate the instructor at the same level as students with little interest in the course. Alternatively, students may take interest level into account when assessing a teacher and therefore not penalize the instructor for their low interest in the course or reward the instructor for high interest.

Both OLS and FE specifications confirm a positive linkage between hard work and teacher effectiveness; but, comparing columns 1-2, we see that the FE coefficient of *Student Effort* is only one-third as large as the OLS coefficient. What this indicates is that more effective teachers (as assessed by students) tend to require greater effort from their students than do less effective teachers. The level of *Student Effort* required in a class varies across instructors, and

---

[22] Based on their preferred cubic specification with faculty fixed effects, and restricting the sample (which comes from the University of California, Santa Barbara) to exclude graduate classes, the model of Bedard and Kuhn predicts that increasing class size from 41 to 148 students would reduce SET ratings by .55.

[23] Numerous student comments were along the following lines: "I thought I would hate this course" or "I don't like economics," but "Professor X did a great job of making the material interesting and relevant."

once we account for this fact, the relationship between student effort and teacher effectiveness weakens, though it does not disappear.

An even greater difference in OLS and FE estimates shows up in the coefficient of *Response Rate*. The reason typically given for including *Response Rate* is student selectivity. Students who are present the day of the course evaluation and fill out the questionnaire may differ from the remaining students. If students who do not complete the questionnaire would tend to rate instructors more harshly than students who do, average SET ratings will be inversely related to the response rate. If the non-respondents would rate the instructor more favorably, the relationship will be positive. Judging from the OLS estimates, it is the students who would rate the instructor highly who tend not to complete evaluation forms, so faculty ratings unfairly suffer when the response rate is low.[24]

When we include faculty effects, the story changes. According to FE estimates, there is no apparent bias from excluding non-responders.[25] Once we account for the individual teacher, there is no relationship between response rate and SET rating. She does not receive higher ratings in her classes that have a high response rate. What the OLS regression picks up is the tendency of more effective teachers to have higher response rates, which is consistent with the findings of Devadoss and Foltz (1996) that class attendance is 9 percentage points higher for instructors who have received teaching awards. Students of effective teachers are more likely to come to class and complete the evaluation forms because they appreciate the teacher—they value what they are getting out of the class. Therefore, in OLS regressions *Response Rate* serves as a

[24] Boex obtained similar results when studying student evaluations of core economics courses.
[25] As Becker and Watts (1999) and Becker and Powers (2001) explain, SET ratings could be influenced by a second type of selectivity if students who drop out would assess their teachers differently than students who remain in the class and if dropout rates vary across instructors. Unfortunately, we do not have the data to control for attrition. But if differences across classes in attrition primarily capture attributes of the individual teacher, as is the case with the response rate, not controlling for faculty attrition should not pose a problem for fixed-effects estimates.

proxy for teacher effectiveness. At least for this university, SET ratings should not be adjusted for the student response rate.

In both specifications, the coefficient of *Expected GPA* is positive but small in value.[26] Based on the estimated value of table 2, column 2, an increase in *Expected GPA* of .37 (the standard deviation for the full sample) would increase the SET rating by .06. According to the estimates of column 1, the increase in SET would be only .03. Thus, there is little evidence that easy grading has an appreciable effect on SET ratings at Kansas State University.

Finally, and somewhat of a surprise, once we control for other characteristics, the coefficient of *Principles* is *positive*. On average, for faculty who teach both principles and non-principles, once we net out for the effect of class size and other variables, students rate principles classes somewhat more highly than non-principles classes. This finding implies that, *if* administrators fully adjusted for variables that are related to SET ratings but not to quality of instructor (notably, class size), instructors would not need to worry that they would pay a price for teaching principles. Of course, the concern articulated by some faculty might be restated as saying that they do not believe that administrators are properly taking such variables as class size into account when evaluating teachers.

Note also that the coefficient of *Principles* is slightly smaller in the FE regression. The implication is that, on average, instructors who teach principles are rated somewhat more highly than instructors who do not teach principles, other things equal. That is, the difference between principles and non-principles ratings is lower when we directly compare instructors who teach

[26] When actual average grade of the class is substituted for expected grade in the regression underlying column 3, its coefficient (t-value) is .197 (3.661), which is in the same ballpark as the estimates of other studies though on the low side. For example, prior estimates (by study) are .15-.25 (Nelson and Lynch, 1984), .32 (Dilts, 1980), .30-.34 (McPherson, 2006), .34-.56 (Krautmann and Sander, 1999), and .53 (Nichols and Soper, 1972). Furthermore, our estimate may overstate the effect of grading policy to the extent that an instructor gives higher grades and is rated more highly in classes that are geared to economics majors or that have better students.

both courses (through FE estimation) than when we allow the coefficient of *Principles* to be influenced by faculty who teach only one course or the other (via OLS). We expand upon this point shortly.

***The Effect of Equitable Grading***

The results to this point have excluded *Grades Equitably*. When we add this variable to the model, we can no longer justify a common structure for the two types of courses.[27] Therefore, we re-estimate the model separately for principles and non-principles samples (columns 4-5, respectively); for purposes of comparison, we also present results for the full sample (column 3).

Because of smaller sample size, especially for principles classes, standard errors are higher for the sub-samples; but the pattern of results is similar to what is reported in column 2. SET ratings are lower for teachers in their first year; pre-semester interest in the courses and response rate are unimportant; and expected GPA of the class has little effect. The estimated coefficients of the two class-size variables are comparable for the two sub-samples and very close to the estimates of column 2. This finding indicates that the results reported in column 2 are not driven by a larger class size for principles classes. To the contrary, for both principles and non-principles classes, SET ratings fall with class size at a comparable (and decreasing) rate. One minor difference in the estimates of columns 4-5 is that greater *Student Effort* appears to improve teacher ratings only in non-principles courses.

For both sub-samples, the effect of equitable grading is large and highly significant. Because the effect appears to be greater in principles classes, we re-estimated the model for the full sample after allowing for a differential effect for the two types of courses. Results are reported in column 6. The coefficient of the interaction term, *Principles*Grades Equitably*, is positive and statistically significant, indicating that principles teachers are rewarded more highly

[27] The assumption of identical coefficients for the principles and non-principles samples is rejected at the .05 level; $F(10, 221) = 2.40$.

than non-principles teachers for increasing grade equity (and penalized more severely for what students perceive as inequitable grading).

Other things equal, when the value of *Grades Equitably* is 3.68, a teacher can expect comparable student ratings in principles and non-principles classes.[28] For *Grades Equitably* = 4.03 (the mean for principles classes), the teacher can expect a SET rating that is .09 higher if teaching principles. When the value of *Grades Equitably* is 4.29 (one standard deviation above the mean for principles classes), the SET premium for principles is .16. Whereas the findings of column 2 suggest that, other things equal, principles classes in general have higher SET ratings, the results of column 6 indicate that principles classes tend to be rated more highly only when equitable grading reaches a threshold (roughly one standard deviation below the mean).

But why is equitable grading rewarded more highly for principles classes? We suggest that a primary reason relates to how grades are determined in the two types of courses. Principles exams are entirely or predominantly multiple-choice in format. Unlike non-principles classes, where a teacher has the opportunity to provide partial credit for wrong answers, grades in principles classes are determined in a more cold and impersonal manner: the computer spits out the scores. Nor do principles classes at this university offer papers or written projects where faculty can provide individual feedback. In this environment, it becomes more difficult to demonstrate greater equity to students. Consistent with this interpretation, both the range and standard deviation of *Grades Equitably* are lower for principles classes.

With the right data, the effect that exam format and non-exam components of the course have on students' perception of grade equity potentially could be examined. Regrettably, such information is missing for the present sample. The primary conclusion of this section is that

[28] According to the estimates of column 6, the SET premium for principles teachers is .262 *Grades Equitably* - .963, which is zero when *Grades Equitably* = 3.68.

equitable grading, as perceived by students, strongly influences their assessment of a teacher. Secondly, the reward for greater equity appears higher in principles classes, perhaps because it is more difficult in such classes to demonstrate increased equity.

## VI. Comparing Faculty Based on their Fixed Effects

Table 3 presents estimates of individual faculty fixed effects for the pooled sample, the principles sample, and the non-principles sample. Estimating the model for the pooled sample forces faculty effects to be the same for principles and non-principles classes. In contrast, estimating the model separately for principles and non-principles samples allows for the possibility that a given faculty member will be rated relatively higher in one type of course than in the other. By comparing estimates from the principles and non-principles samples, we can estimate the extent to which teacher effectiveness (as judged by students) carries over across courses, as opposed to the situation where faculty have a strong comparative advantage in one type of course.

The first observation drawn from the table is that differences in teaching effectiveness are substantial. For the pooled sample, the difference in estimated coefficients of the top-ranked and lowest-ranked faculty is .89. For the principles and non-principles samples, the differences are 1.02 and .88, respectively. These numbers are in the same ballpark as those obtained by McPherson (2006) for economics faculty at the University of North Texas (1.44 for principles and .90 for non-principles).[29] Student evaluations provide only one dimension of teaching effectiveness, and they should be supplemented with other measures; but differences in SET

[29] Based on a different dependent variable, end-of-semester test score, Watts and Bosshardt (1991) also found evidence of substantial faculty fixed effects. For both survey and principles courses, they estimated that differences in the test scores of the most effective and least effective teachers amounted to at least 20 percent of the points possible on the test.

ratings after accounting for other factors are sufficiently large to provide a basis for comparison.[30]

An inspection of the coefficient estimates of table 3 reveals a second finding: Faculty who are rated as good (weak) principles teachers are also viewed as good (weak) non-principles teachers. For the ten faculty who taught both principles and non-principles courses during the sample period, the correlation between the coefficients of columns 2 and 3 is .843. Decisions on faculty teaching assignments would be easier if the faculty who were weak teachers in one category of courses were strong in the other category, but that does not appear to be the case.

A third observation is that faculty who teach principles are, in general, rated more highly than faculty who do not teach principles, a point previously made when comparing the coefficient of *Principles* in OLS and FE specifications. Based on the estimates of column 1, five of the seven lowest rated faculty exclusively taught non-principles courses during the sample period.[31] Of the seven highest rated faculty, all seven taught principles. Thus, faculty who do not teach principles are more likely to be on the low end of the distribution and less likely to be at the top.

We want to be careful to point out that some faculty who do not teach principles receive above average student evaluations, and some principles teachers are rated poorly. Also, this study is limited to undergraduate instruction, and it is possible that faculty who do not teach principles do well in the graduate courses they teach. Finally, we make no attempt to generalize these results beyond the particular university studied. What we do say, to those in the public who argue that the best undergraduate teachers should be assigned to large introductory courses,

---

[30] Pallett (2006, p. 57) cautions against making too much of minor differences ("Is there really a difference between the student ratings averages of 4.0 and 4.1?"). Accordingly, he recommends classifying faculty teaching on the basis of no more than five discrete categories, such as "outstanding," "exceeds expectations," meets expectations," "needs improvement but making progress," and "fails to meet expectations."

[31] Included in the list is faculty number 5, who serves as the reference category for non-principles courses.

is that, based on the SET data we analyze, there is evidence that this is occurring in the economics department we studied.

## VII. SUMMARY AND CONCLUSIONS

We study student evaluations of economics faculty at Kansas State University over 19 semesters to compare evaluations in principles and non-principles courses. Although we present OLS results for comparison, we rely on estimates that allow for faculty fixed effects. Fixed-effects estimates offer several advantages over ordinary least squares. First, they control for unobserved, time-invariant faculty characteristics that are correlated with explanatory variables and bias OLS estimates. Second, they allow us to account for non-random teaching assignments. If the best teachers teach principles, which is what we find, the coefficient of *Principles* is positively biased when estimated by ordinary least squares. Finally, fixed-effects estimates provide a basis for comparing the relative teaching effectiveness of faculty.

Unadjusted SET ratings are, on average, .3 point lower in principles classes than non-principles classes (on a 4.0 scale). But once we control for other variables, the coefficient of *Principles* turns positive. Evaluations are not inherently lower in principles, despite the large number of students who take the course because it is required for their major. Based on differences in just one variable, average class size, our model predicts that raw SET ratings will be approximately .5 lower in principles. These findings suggest that it is inappropriate to compare unadjusted student assessments, but they also indicate that departments can potentially adjust for factors, such as class size, that influence these assessments.

Other things equal, faculty receive lower evaluations their first year at the university. Thereafter, there is no evidence that years of teaching experience influence SET ratings. Faculty appear to learn quickly how to raise student assessments, and they make the necessary adjustments.

Class GPA is not an important determinant of student ratings. The coefficient of expected class GPA, while positive, is consistently small in value. In contrast, the coefficient of *Grades Equitably* is large and highly significant. Our interpretation is that students place a high value on fair treatment. Ultimately, the distribution of grades—higher grades for students who learn more or study harder—is more important than the average grade in the class. The rewards for equitable grading appear larger for principles classes, consistent with evidence that it is more difficult to raise the equity rating in principles classes. We hypothesize that impersonal nature of principles grades, determined predominately by multiple-choice exams (for which partial credit and helpful feedback are more difficult), contribute to the challenge faced by principles teachers.

SET ratings are positively related to student effort, but the quantitative relationship between the two variables falls by two-thirds once we account for faculty effects. This finding points to the fact that more effective teachers require more effort from their students than less effective teachers. Student assessments do not depend on reported pre-semester interest in the course. Either students do not punish/reward an instructor for their interest in the course, or ex post interest in the course is only weakly correlated with a priori interest.

When estimated by ordinary least squares, the relationship between SET and response rate is positive, suggesting that a low response rate leads to an underassessment of a faculty member's teaching effectiveness. But once faculty effects are included, the response rate is irrelevant. More effective teachers have higher response rates, consistent with prior research that class attendance is higher for better teachers. The positive coefficient of response rate in OLS estimates captures the influence of unmeasured characteristics of the teacher. At least for this university, OLS estimates should not be adjusted for response rate. To do so would penalize better teachers for the higher response of their students.

Faculty effects are important quantitatively and statistically. Other things equal, the top-rated faculty member in the department can expect a SET rating 0.9 point higher than the lowest rated faculty member. Faculty effects obtained separately for principles and non-principles courses are highly correlated—faculty who are rated as good (weak) principles teachers tend to be rated as good (weak) non-principles teachers.

The top-rated teachers taught principles during the sample period; most of the teachers with low faculty effects did not. Thus, at the public university that was the focus of this study, the most highly rated teachers teach the courses with the largest class size: principles of macroeconomics and principles of microeconomics. Whether these results generalize to other colleges and universities is an open question. Further research is needed if we are to fully understand the relationship between the student assessments of principles and non-principles courses and the assignment of teachers to these courses.

# VIII. Figures and Tables

**Table 3.1 Sample Statistics for Principles and Non-Principles Samples**

| Variable | Mean | Std. Dev. | Min. | Max. |
|---|---|---|---|---|
| I. Principles (N = 94) | | | | |
| *Teacher Effectiveness* | 3.609 | 0.407 | 2.500 | 4.500 |
| *Year 1* | 0.170 | 0.378 | 0.000 | 1.000 |
| *Year 2* | 0.160 | 0.368 | 0.000 | 1.000 |
| *Year 3-4* | 0.191 | 0.396 | 0.000 | 1.000 |
| *Class Size/10* | 14.786 | 2.025 | 8.000 | 17.500 |
| *Class Size$^2$/1000* | 0.223 | 0.053 | 0.064 | 0.306 |
| *Student Interest* | 2.886 | 0.147 | 2.570 | 3.250 |
| *Student Effort* | 3.641 | 0.155 | 3.300 | 4.030 |
| *Response Rate* | 0.561 | 0.130 | 0.209 | 0.789 |
| *GPA* | 2.304 | 0.185 | 1.850 | 2.950 |
| *Expected GPA* | 2.352 | 0.239 | 1.850 | 3.200 |
| *Grades Equitably* | 4.034 | 0.257 | 3.100 | 4.500 |
| II. Non-Principles (N = 190) | | | | |
| *Teacher Effectiveness* | 3.906 | 0.494 | 2.660 | 5.000 |
| *Year 1* | 0.053 | 0.224 | 0.000 | 1.000 |
| *Year 2* | 0.074 | 0.262 | 0.000 | 1.000 |
| *Year 3-4* | 0.068 | 0.253 | 0.000 | 1.000 |
| *Class Size/10* | 4.063 | 2.188 | 0.300 | 11.500 |
| *Class Size$^2$/1000* | 0.021 | 0.022 | 0.000 | 0.132 |
| *Student Interest* | 3.352 | 0.407 | 2.400 | 4.400 |
| *Student Effort* | 3.753 | 0.256 | 3.000 | 5.000 |
| *Response Rate* | 0.767 | 0.116 | 0.462 | 1.000 |
| *GPA* | 2.738 | 0.365 | 2.100 | 4.000 |
| *Expected GPA* | 2.786 | 0.346 | 2.060 | 4.000 |
| *Grades Equitably* | 4.048 | 0.398 | 2.900 | 5.000 |

**Table 3.2 Estimated Determinants of SET Ratings**

| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| *Year 1* | -.355*** | -.411*** | -.245*** | -.194 | -.286** | -.201** |
| | (4.73) | (3.67) | (2.79) | (1.36) | (2.36) | (2.27) |
| *Year 2* | -.048 | -.140 | -.131 | -.155 | -.174 | -.124 |
| | (.66) | (1.31) | (1.59) | (1.33) | (1.53) | (1.51) |
| *Year 3-4* | .002 | -.123 | -.104 | -.131 | -.099 | -.099 |
| | (.03) | (1.19) | (1.31) | (1.13) | (.93) | (1.25) |
| *Class Size/10* | -.084*** | -.118*** | -.084*** | -.106 | -.101*** | -.088*** |
| | (3.93) | (6.32) | (5.78) | (.86) | (3.16) | (6.05) |
| *Class Size$^2$/1000* | .030** | .038*** | .027*** | .038 | .034 | .029*** |
| | (2.47) | (3.74) | (3.37) | (.79) | (1.12) | (3.68) |
| *Student Interest* | .052 | .069 | .063 | .192 | .056 | .083* |
| | (.75) | (1.10) | (1.29) | (1.10) | (.95) | (1.71) |
| *Student Effort* | .858*** | .314*** | .186** | -.040 | .175** | .187** |
| | (8.13) | (3.28) | (2.50) | (.23) | (2.05) | (2.53) |
| *Response Rate* | .883*** | .096 | .157 | -.029 | .110 | .130 |
| | (4.69) | (0.57) | (1.19) | (.12) | (.62) | (1.00) |
| *Expected GPA* | .088 | .155** | .078 | .171 | .074 | .082* |
| | (1.24) | (2.49) | (1.60) | (1.57) | (1.18) | (1.70) |
| *Grades Equitably* | | | .606*** | .842*** | .588*** | .565*** |
| | | | (12.60) | (6.70) | (10.17) | (11.12) |
| *Principles* | .370*** | .242** | .092 | | | -.963** |
| | (2.78) | (2.13) | (1.04) | | | (2.07) |
| *Principles*Equity* | | | | | | .262** |
| | | | | | | (2.31) |
| Technique/Sample | OLS/Full | FE/Full | FE/Full | FE/Prin | FE/Non-P | FE/Full |
| $R^2$ (adjusted) | .530 | .735 | .842 | .855 | .841 | .845 |

Notes: Regressions also control for semester. Sample size is 284 for full sample, 94 for principles (column 4), and 190 for non-principles (column 5). Numbers in parentheses are absolute values of t-statistics, and asterisks indicate significance at the .10, .05, and .01 level (two-tailed test).

**Table 3.3 (a) Estimated Faculty Fixed Effects**

| Faculty Number | (1) Full Sample | (2) Principles | (3) Non-Principles |
|---|---|---|---|
| 1 | 0.207 | 0.220 | 0.189 |
| | (1.464) | (1.109) | (1.532) |
| 2 | 0.228* | 0.014 | 0.325*** |
| | (1.932) | (0.088) | (4.190) |
| 3 | 0.183 | | 0.285** |
| | (1.266) | | (2.603) |
| 4 | 0.462*** | 0.373*** | 0.566*** |
| | (4.108) | (2.898) | (7.169) |
| 5 | -0.031 | | |
| | (0.255) | | |
| 6 | -0.028 | | 0.085 |
| | (0.117) | | (0.366) |
| 7 | 0.335** | 0.340** | 0.330*** |
| | (2.581) | (2.120) | (2.893) |
| 8 | 0.378*** | 0.323** | 0.482*** |
| | (3.157) | (2.132) | (5.201) |
| 9 | 0.310 | | 0.319 |
| | (1.286) | | (1.413) |
| 10 | 0.371*** | | 0.436*** |
| | (3.047) | | (5.859) |
| 11 | 0.534*** | | 0.578*** |
| | (4.157) | | (6.652) |
| 12 | 0.580*** | 0.716*** | |
| | (3.335) | (4.065) | |
| 13 | 0.710*** | 0.688*** | |
| | (6.138) | (5.012) | |
| 14 | | | |

**Table 3.3 (b) Estimated Faculty Fixed Effects**

| Faculty Number | (1) Full Sample | (2) Principles | (3) Non-Principles |
|---|---|---|---|
| 15 | 0.021 | | 0.119 |
| | (0.173) | | (0.772) |
| 16 | 0.104 | -0.059 | 0.229** |
| | (0.945) | (0.404) | (2.345) |
| 17 | 0.621*** | 0.709*** | 0.561*** |
| | (5.114) | (4.385) | (3.139) |
| 18 | 0.427*** | | 0.477*** |
| | (3.510) | | (3.328) |
| 19 | 0.576*** | 0.597*** | |
| | (5.513) | (5.515) | |
| 20 | 0.163 | | 0.250* |
| | (1.408) | | (1.936) |
| 21 | 0.254* | 0.210 | |
| | (1.714) | (1.104) | |
| 22 | 0.811*** | 0.793*** | 0.881*** |
| | (6.923) | (5.659) | (4.420) |
| 23 | 0.860*** | 0.964*** | 0.820*** |
| | (7.593) | (6.339) | (5.013) |
| 24 | 0.589*** | 0.690*** | 0.465*** |
| | (5.625) | (5.060) | (3.807) |

Notes: Estimates are based on the regression underlying table 2, column 3, and the same equation estimated separately for the principles and non-principles samples. The reference category is faculty #14 for the pooled and principles samples and #5 for the non-principles sample. Numbers in parentheses are absolute values of t-statistics.

# References

Aigner, D., and Thum, F. 1986. “On student evaluation of teaching ability” *Journal of Economic Education* 17(4): 243-65.

Akin, J. 1992. “A child health production function estimated from longitudinal data.” *Journal of Development Economics* 38(2): 241-72.

Anderson, J., Johnstone B., and Remley D. 1999. “Breast-feeding and cognitive development: a meta-analysis.” *The American Journal of Clinical Nutrition* 70(4): 525-35.

Anderson, P., Butcher, K., and Levine, P. 2003. “Maternal employment and overweight children.” *Journal of Health Economics* 22 (3): 477-504.

Angelsen, N., Vik, T., Jacobsen G., and Bakketeig L. 2001. “Breast feeding and cognitive development at age 1 and 5 years.” *Archives of disease in childhood*. 85(.): 183-88.

Angrist, J., and Lavy, V. 1996. “The effects of teen childbearing and single parenthood on childhood disabilities and progress in school.” NBER Working Paper 5807.

Barrera, A. 1990. “The interactive effects of mother's schooling and unsupplemented breastfeeding on child health.” *Journal of Development Economics* 34(1): 81-98.

Becker, W., and Powers, J. 2001. “Student performance, attrition, and class size given missing student data.” *Economics of Education Review* 20(4): 377-88.

Becker, W., and Watts, M. 1996. “Chalk and talk: a national survey on teaching undergraduates.” *American Economic Review Papers and Proceedings* 86(2): 448-53.

Becker, W., and Watts, M. 1999. “How departments of economics evaluate teaching.” *American Economic Review* 89 (2): 344-49.

Bedard, K., and Kuhn, P. 2005. “Where class size really matters: class size and student ratings of instructor effectiveness.” working paper, University of California, Santa Barbara.

Berger L., Hill J., and Waldfogel J. 2005. “Maternity leave, early maternal employment and child health and development in the US.” *The Economic Journal* 115(501): F29-F47.

Blanden, J., Gregg, P., Macmillan, L. 2006. “Explaining intergenerational income persistence: non-cognitive skills, ability and education.” Working Paper.

Blau, David. 1991. “Child care demand and labor supply of young mothers over time.” *Demography* 28(3): 333-51.

Blau, D., and Robins, P. 1991. “Turnover in child care arrangements.” *The Review of Economics and Statistics* 73(1): 152-7.

Boex, L. 2000. “Attributes of effective economics instructors: an analysis of student evaluations” *Journal of Economic Education* 31(3): 211-27.

Booth A., and Kee, H. 2006. “Birth order matters: The effect of family size and birth order on educational attainment.” CEPR Discussion Paper 5453.

Bosshardt, W., and Watts, M. 2001. “Comparing student and instructor evaluations of teaching” *Journal of Economic Education* 32 (1): 3-17.

Case, A., Fertig, A., and Paxson, C. 2005. “The lasting impact of childhood health and circumstance.” *Journal of Health Economics* 24(2): 365-89.

Cebu Study Team. 1992. “A child health production function estimated from longitudinal data.” *Journal of Development Economics* 38(2): 323-51.

Centra, J. (1978). “Using student assessments to improve performance and vitality” in W. R. Kirschling (ed.), *Evaluating Faculty Performance and Vitality*. New Directions for Institutional Research 20, San Francisco: Jossey-Bass.

Clifford, T. 2003. “Breastfeeding and obesity.” *British Medical Journal* 327(7420): 879-80.

Connor, S. 2003. “Scientists link obesity to `thrifty gene' of our ancestors.” *The (London) Independent*. 2003-02-07. Retrieved on 2008-03-22.

Devadoss, S., and Foltz, J. “Evaluation of factors influencing student class attendance and performance” *American Journal of Agricultural Economics* 78(3): 499-507.

Dietz, William. 1991. “Factors associated with childhood obesity.” *Nutrition* 7(4): 290-1.

Dilts, D. 1980. “A statistical interpretation of student evaluation feedback” *Journal of Economic Education* 11(2): 10-5.

Feldman, K. (1983). “Seniority and experience of college teachers as related to evaluations they receive from student.” *Research in Higher Education* 18(1): 3-124.

Finegan, T., and Siegfried, J. “Are student ratings of teaching effectiveness influenced by instructors’ English language proficiency?” *American Economist* 44(2): 17-29.

Fertig, A., Glomm, G., and Tchernis, R. 2006. “The connection between maternal employment and childhood obesity: Inspecting the mechanisms.” Center for Applied Economics and Policy Research, Economics Department, Indiana University Bloomington, Caepr Working Papers.

Finkelstein, E., Ruhm, C., and Kosa, K. 2005. "Economic Causes and Consequences of Obesity." *Annual Review of Public Health* 26(1): 239-57.

Gillman, M., Rifas-Shiman, S., Camargo, C., Berkey, C., Frazier, A., Rockett, H., Field, A., Colditz, G. 2001. "Risk of overweight among adolescents who were breastfed as infants." *JAMA: The Journal of the American Medical Association* 285(19): 2461-7.

Grummer-Strawn, L. 2004. "Does breastfeeding protect against pediatric overweight? analysis of longitudinal data from the centers for disease control and prevention pediatric nutrition surveillance system." *Pediatrics* 113(2): 81-6.

Hastie, T., and Tibshirani, R. 1990. Generalized Additive Models. Chapman and Hall/CRC: USA.

Hausman, J. 1978. "Specification tests in econometrics." *Econometrica* 46(6): 1251-71.

Heckman, J. 1979. "Sample selection bias as a specification error." *Econometrica* 47(1): 153-61.

Hediger, Mary L., Mary D. Overpeck, Robert J. Kuczmarski, and W. June Ruan. 2001. "Association between infant breastfeeding and overweight in young children." *JAMA: The Journal of the American Medical Association* 285, (19) (May 16): 2453-60.

Heitmueller, A. 2007. "The chicken or the egg?: Endogeneity in labour market participation of informal careers in England." *Journal of Health Economics,* 26(3): 536-59.

Horwood, L., and Fergusson D. 1998. "Breastfeeding and later cognitive and academic outcomes." *Pediatrics* 101(1): 1-6.

Isely, P., and Singh, H. 2005. "Do higher grades lead to favorable student evaluations?" *Journal of Economic Education* 36 (1): 29-42.

Jacobson, S., Chiodo, L., and Jacobson, J. 1999. "Breastfeeding effects on intelligence quotient in 4- and 11- year old children." *Pediatrics* 103(1): 1-6.

Jacobson S., and Jacobson J. 2002. "Breastfeeding and IQ: evaluation of the socio-environmental confounders." *Acta Pediatrica*; 91(3): 267-74.

Kalmijn, M., and Kraaykamp, G. 2005. "Late or Later? A sibling analysis of the effect of maternal age on children's schooling." *Social Science Research* 34(3): 634-50.

Kurinu N., Shiono P., Ezrine S., and Rhoads, G. 1989. "Does maternal employment affect breast-feeding?" *American Journal of Public Health* 79(.): 1247-50.

Li, L., Parsons, T. and Power, C. 2003. "Breast feeding and obesity in childhood: Cross sectional study." *British Medical Journal* 327(7420): 904-5.

Liese, A. 2001. "Inverse association of overweight and breast feeding in 9 to 10-year-old children in Germany." *International Journal of Obesity* 25(11): 1644-50.

Liu, E. 2006. "Maternal full-time employment and childhood obesity." University of Southern California. Working Paper.

Locard, E. 1992. "Risk factors of obesity in a five year old population. parental versus environmental factors." *International Journal of Obesity* 16(10): 721-9.

Malloy, M, and Berendes, H. 1998. "Does breast-feeding influence intelligence quotients at 9 and 10 years of age?" *Early Human Development* 50(2): 209-17.

Malloy, M., Rhoads G., Schramm W., and Land G. 1989. "Increasing cesarean section rates in very low-birth weight infants. Effect on outcome." *JAMA: The Journal of the American Medical Association* 262(11): 1475-8.

Marsh, H., and Duncan, M. 1992. “Students’ evaluations of university teaching: a multidimensional perspective” in J. C. Smart (ed.), Higher Education: Handbook of Theory and Research 8, New York: Agathon Press.

Matthews, K. 1998. “Maternal infant-feeding decisions: Reasons and influences.” *The Canadian Journal of Nursing Research* 30(2): 177-98.

McPherson, M. 2006. “Determinants of how students evaluate teachers” *Journal of Economic Education* 37(1): 3-20.

Mei, Z. 2002. “Validity of body mass index compared with other body-composition screening indexes for the assessment of body fatness in children and adolescents.” *The American Journal of Clinical Nutrition* 75(6): 978-85.

Michaelsen, K., Lauritzen, L., Jorgensen, M., and Mortensen, E. 2003. “Breast-feeding and brain development.” *Scandinavian Journal of Nutrition* 47(3): 147-151.

Miller, P., Mulvey, C., and Martin, N. 2005. “Birth weight and schooling and earnings: estimates from a sample of twins.” *Economic Letters* 86(3): 387-92.

Mirus, R. 1973. “Some implications of student evaluations of teachers.” *Journal of Economic Education* 5(1): 35-37.

Mortensen E., Michaelsen K., Sanders S., and Reinisch J. 2002. “The association between duration of breastfeeding and adult intelligence.” *The Journal of the American Medical Association* 287(18): 2365-71.

Nelson, J., and Lynch, K. 1984. “Grade inflation, real income, simultaneity, and teaching evaluations.” *Journal of Economic Education* 15(1): 21-37.

Nichols, A., and Soper, J. 1972. “Economic man in the classroom.” *Journal of Political Economy* 80(5): 1079-83.

Pallett, W. 2006. “Uses and abuses of student ratings.” in P. Seldin & Associates, *Evaluating faculty performance,* Bolton, MA: Anker: 50-65.

Paracchini, V. 2005. “Genetics of leptin and obesity: A HuGE review.” *American Journal of Epidemiology* 162(2):101-14.

Plug, E., and Vijverberg, W. 2005. “Does family income matter for schooling outcomes? Using adoptees as a natural experiment.” *The Economic Journal* 115(506): 879-906.

Powers, E. 2003. “Children's health and maternal work activity: Estimates under alternative disability definitions.” *The Journal of Human Resources* 38(3): 522-56.

Raj, V. 1998. “The role of social support in breastfeeding promotion: A literature review.” *Journal of Human Lactation* 14(1): 41-45.

Roe, B., Whittington, L., Fein, S., Teisl, M. 1999. “Is there competition between breast-feeding and maternal employment?” *Demography* 36(2): 157-71.

Ruhm, C. 2005. “Healthy living in hard times.” *Journal of Health Economics* 24(2): 341-63.

Saunders, K. (2001). “The influence of instructor native language on student learning and instructor ratings.” *Eastern Economic Journal* 27(3): 345-53.

Schimek, M. 2000. “Smoothing and regression: Approaches, computation and application.” Wiley: New York.

Seiver, D. 1983. "Evaluations and grades: a simultaneous framework" *Journal of Economic Education* 14(3): 32-8.

Senauer, B., and Kassouf, A. 2000. "The effects of breastfeeding on health and the demand for medical assistance among children in Brazil." *Economic Development and Cultural Change* 48(4): 719-36.

Silva, A. 1998. "Trends in low birth weight: A comparison of two birth cohorts separated by a 15-year interval in Ribeirão Preto, Brazil." *Bulletin of the World Health Organization* 76(1): 73-84.

Smith, J., and Ingham L. 2001. "Breastfeeding and the measurement of economic progress." *Journal of Australian Political Economy* 47(.): 51-72.

Staiger, D. and Stock, J. 1994. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3): 557-86.

Story, M., Kaphingst, K. and French, S. 2006. "The role of child care settings in obesity prevention." *The Future of Children* 16(1): 143-68.

Strauss, R. 1999. "Influence of the home environment on the development of obesity in children." *Pediatrics* 103(6): e85.

Sullivan M., Sonya, J., and Michele, A. 2004. "Family characteristics associated with duration of breastfeeding during early infancy among primiparas." *Journal of Human Lactation* 20(2): 196-205.

Vandewater, E., Shim, M., Caplovitz, A. 2004. "Linking obesity and activity level with children's television and video game use." *Journal of Adolescence* 27(1): 71-85.

Von Kries, R. et al. 1999. ”Breast feeding and obesity: Cross sectional study.” *British Medical Journal* 319(7203): 147- 50.

Watts, M., and Bosshardt, W. 1991. “How instructors make a difference: panel data estimates from principles of economics courses.” *Review of Economics and Statistics* 73(2): 336-40.

Weinberg, B., Fleisher, B., and Hashimoto, M. 2007. “Evaluating methods for evaluating instruction: the case of higher education.” NBER Working Paper 12844.

Whitaker, R. 1997. “Predicting obesity in young adulthood from childhood and parental obesity.” *New England Journal of Medicine* 337(13): 869-73.

Wosje, K. 2004. “Lactation, weaning, and calcium supplementation: Effects on body composition in postpartum women.” *The American Journal of Clinical Nutrition* 80(2): 523-8.