FLORIDA STATE UNIVERSITY

COLLEGE OF SOCIAL SCIENCES

INVESTIGATING ROBUSTNESS IN THE ENGEL AND ROTHBARTH MODELS: AN EMPIRICAL ESTIMATION STUDY

By

M. PAULA ARCE-TRIGATTI

A Thesis submitted to the
Department of Economics
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Spring Semester, 2009

The members of the Committee approve the Thesis of Paula Arce-Trigatti defended on August 14th, 2008,

____
David Macpherson
Professor Directing

____
Thomas McCaleb
Committee Member

____
Stefan Norrbin
Committee Member

The Office of Graduate Studies has verified and approved the above named committee members.

ACKNOWLEDGEMENTS

I would like to first acknowledge the help of my mother, Beatriz Trigatti, whose expertise in programming and mathematics has been an invaluable source of assistance throughout every step of this project. Secondly, I would like to acknowledge the helpful discussions with a veteran in the paper writing field, my father, Dr. Pedro Arce, whose encouraging words always shed light in times of trouble. Thank you tremendously, to both of you.

Next I would like to acknowledge Daniel Betancourt, whose insistence on discussing theoretical problems until comprehension inevitably led to a deeper understanding of the subject matter, as well as to a plethora of more questions. Thank you for being my 'personal trainer.'

Dr. Norrbin, who initially welcomed me into the Master's Program at FSU and later, continually supported my efforts as a student within the program – thank you for taking the time to help me succeed.

My thesis director, Dr. Macpherson: I would like to acknowledge his time and patience in helping me learn during this experience. His easy-going and encouraging demeanor is uniquely special in a professor of his caliber and I feel very lucky to have had the opportunity to work with him. Thank you enormously for your guidance and support.

And finally, my sister, Andrea, and friends who have been supportive and encouraging in many, many respects. Thank you for helping me see this through to the end.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Currently, there exist no viable data in order to accomplish the answering of the question: how much does a child cost? As a result, several models have been proposed and engineered to handle the approach to this problem in what is seen in each as the most sound way possible. The two frontrunners that have found the greatest popularity are the Engel Model and the Rothbarth Model, both of which have been developed on the idea that an appropriate proxy for the standard of living of a household should be able to assist in identifying equally well-off households; once these connections have been made, the pathway for determining the amount necessary to put a household containing children back to the same standard of living enjoyed previous to the presence of children is ultimately cleared. This required amount is what is deemed as the 'cost' of a child. Despite the similarities in the underlying assumptions and structure of either of these two models, it has been shown that results can greatly vary. Affected and interested parties include those planning to have children, statewide child support guidelines that use these figures to calculate award amounts, and most importantly, the children themselves. Thus, studies that attempt to investigate the precision of these popular models are quite necessary. While accuracy in and of itself cannot be compared with a true standard, instead, model robustness can be tested, which is the goal of this study. Findings from this study indicate that the Rothbarth Model is perhaps more sensitive to the underlying data set than the Engel Model, which suggests that the Engel Model may be a more trustworthy choice in this particular duel. Additionally, a second result from this study suggests that the Engel Model may be quite sensitive to model specification. This result is potentially the most relevant at the law-making level, where truthfulness in the estimation of child costs has the greatest impact on children.

1. INTRODUCTION

The initial effort of the U.S. government to quell disparities and difficulties in producing child support guidelines within the states arrived via the debut of the Child Support Enforcement Amendments of 1984. Until that time, no uniform precedent had been established or accepted; as a result, many states relied upon the individual efforts of judges and localized law firms to author supposedly unbiased and equitable guidelines on a case by case basis (Williams, 1987). Perhaps the most obvious consequence of adopting such a method is the certain resulting lack of uniformity amongst the child support awards, since, essentially, this approach is founded on localized opinion rather than on (a somewhat more sound) numerical and economics-supported backing. Indeed, as some studies have pointed out (see White and Stone, 1976; Yee, 1979; Garfinkel, 1992), the awards tended to be inequitable and leaned towards the lower end of the numeric spectrum. Moreover, one can also imagine the extra administrative inefficiencies (see, for example, Pirog-Good and Brown, 1996) created by having to wait for the results of specific cases to be considered, rather than enabling the associated parties to easily access the prescribed amounts obtained from a predetermined formula or model.

As such, sections 15 and 18 of the 1984 amendments made it mandatory for states to appoint a committee to develop a numerical formula for setting child support guidelines. Despite the requirement that the formula have numeric and descriptive criteria as a foundation, the amendment did *not* require the actual implementation of this formula; an advisory action, rather than a presumptive one, was imposed[1]. It was not until the passing of the Family Support Act of 1988 that a presumptive action was prescribed. Indeed, the Act made it mandatory for judges to utilize the guidelines set forth by their respective states in deciding child support award amounts (unless written justification is provided for the deviation from its use), but it did not, however, mandate or sponsor a specific mathematical formula or model[2]. This notable exclusion in the recommendations has fostered what is essentially an underground war in whose model is, in fact, the most

[1] Public Law 98-378, Sections 15 and 18.
[2] Public Law 100-485.

appropriate choice. While each state has the option to adopt the resulting model it has determined to be the most just and fair, there are two models that have gained the most popularity in this battle: the model commonly known as the 'Engel Model' and secondly, what has now become the 'Rothbarth Model'.

Despite the positive and renewed energies surely associated with a more ordered approach to child support guidelines, the two models introduced above serve as a platform for the following bleak realization: while a scenario in which one can easily refer to a predetermined schedule of payments given a certain level of parental income sounds extraordinary, in reality, it is extremely difficult to materialize. The underlying reason for such an austere viewpoint is the inherent obstacles involved in estimating how much a child is worth. For example, given a household, it is not necessarily clear who consumes what (imagine an attempt to divide electricity usages among a two-parent, two-child household), there exists no data source that conveys each and every cost incurred due to the presence of a child, and furthermore, there may be economies of scale present under certain consumption choices that are impossible to pinpoint. As a result, any effort directed at explicitly determining the true cost of a child would ultimately falter due to the aforementioned reasons.

However, what is an impossible feat for some may be a mere intriguing challenge for others. Thomas Espenshade made his debut in the realm of estimating the cost of children with his initial contribution, preceding the 1984 amendments and foreshadowing the importance of this topic in the national lawmaking arenas, *The Cost of Children in Urban United States*, in 1973. A superiorly organized literature review of studies prior to his that delved into answering the very question of how much a child cost provided a solid foundation for what was to be Espenshade's next pertinent contribution to the field, a 1984 book entitled, *Investing in Children*. Indeed, the principal economic research initially consulted for most of the child support guideline recommendations was taken from Espenshade's 1984 study, which included estimates of expenditures on children based on data from the 1972-1973 Consumer Expenditure Survey (henceforth, "CEX") provided by the Bureau of Labor Statistics (Morgan, 2004). Espenshade approached the problem of estimating these expenditures by utilizing what is commonly known as the Engel Model, which will be discussed in further detail in Section 3.1 of this study. Thus, Espenshade was able to, in some sense, answer the question of what a child costs, and in the process, produce estimates that were later incorporated (by adjusting Espenshade's figures into 1986 price levels) into the government-sponsored report put together by Robert Williams for the U.S. Department of Health and Human Services[3].

As one can see, the final product of the child support guidelines requirement is actually the result of an estimation technique rooted in economics; indeed, there are several child cost estimation techniques presented in the academic literature in addition to the Engel method (for examples and criticisms, see Section 2 of this paper). For example, yet another approach to estimation was proposed by Erwin Rothbarth (1943), which has also become quite a popular method, and will be further discussed in Section 3.2 of this study. Perhaps the primary promoter of the Rothbarth Model is David Betson, currently a professor at the University of Notre Dame, who has also made quite an impact with his efforts at answering the very same question presented to Espenshade. Betson was commissioned in a government-sponsored effort in 1990 to produce a report that considered five alternative approaches to estimating expenditures on children, among them in the paper, the Engel and Rothbarth approaches were considered. The comprehensive review of these methods was based on data from the 1980 – 1986 CEX. After his experience with these models, providing updates to the Engel and Rothbarth Model results using more current CEX-based data in 2000 and 2006, Betson (2006) arrived at the conclusion that the Engel model was indeed the weaker of the two, because of the heroic assumptions needed, and that the more reliable model of the two was the Rothbarth Model (see Section 2.2 for a more detailed criticism of this topic). As far as is known, no other study in the literature has since attempted to update either of these estimation results, nor has any study attempted to replicate Betson's findings from the Engel Model in his 2000 study or the Rothbarth Model in his 2006 study.

A third major contributor to these estimations comes from the U.S. Department of Agriculture, which has been producing estimates of child expenditures since the 1960s, and will be further discussed in the following section.

[3] (see Robert G. Williams, Development of Guidelines for Child Support Orders: Advisory Panel Recommendations and Final Report (U.S. Dept of HHS, Office of Child Support Enforcement, 1987).

The nature of this topic, that is, the business of answering just what the costs are associated with a child, is a politically sensitive one – several relevant parties are directly affected by the results of these studies, including the future financial well-being of either parent, taxpayer dollars (due to government sponsored programs that attempt to handle parental failure to pay child support, for example), and more importantly, the children themselves. That being the case, any efforts towards a solution to this question should be approached with a superior level of meticulous caution. Because it appears as though there is an insufficient amount of empirical cross-checking of proposed models, as alluded to previously, the main efforts undertaken in this thesis are attempts to help remedy this situation.

Firstly, this study seeks to replicate the Rothbarth results from Betson's 2006 study, as well as the Engel results from Betson's 2000 study, using a sub-sample of the CEX data spanning a similar time period used in Betson's studies. While this is not the ultimate goal of this thesis, it should facilitate a discussion for what is the true intention behind this body of work: testing for model robustness. Moreover, what is meant by model robustness in this instance is a notable indifference to the data set employed when running the model. That is, results that are consistent and invariant to the underlying data set shall be considered a strong indication of the level of accuracy of the model being tested.

Accordingly, testing for model robustness will be achieved in the following manner: after implementing the deletions as noted in the various Betson studies (see Section 5.2 for a more detailed explanation), a further data set restriction will be imposed, one that is not necessarily economically motivated. That is, only those respondents to the CEX that have completed at least three of the four interview attempts will be included in a separate data set, eliminating almost half of the observations from the initial sub-sample. Thus, two data sets will remain: the initial sub-sample, and a further restricted data set that contains about half of the initial sub-sample's observations. It is then assumed that applying each data set to both the Engel and Rothbarth Models will produce results that will assist in the discussion on model robustness.

A further examination into the Engel Model will then be presented, in an effort to address some of the concerns surrounding the credibility of this approach. An Engel-based approach that was first introduced in a report to the Florida Legislature by the Department of Economics at Florida State University (2004) for consideration in the mandatory updates to state child support laws will be tested using the two data sets developed for the original Engel and Rothbarth Models. This model is essentially a simplified version of the more detailed Engel Model employed by Betson (2000) (see Section 4.2 for a detailed explanation of differences), and will henceforth be referred to as the "Age-Invariant" Model.

A third consequence of this thesis, then, will be translating the regression results obtained in the previous step into an appropriate answer to the question first posed in the abstract of this work: what is the cost of a child? Actual costs will not be presented; rather, the more common presentation of shares of total expenditures dedicated to a child will be presented instead. A comparison of these share estimations will be discussed to gain a better understanding of the performance of each model, as well as to contribute an empirical estimation into the cost of children to the literature.

The remainder of this paper will proceed as follows: we will begin with a literature review of current and past findings in Section 2. In Section 3, there will be a theoretical discussion behind the Engel and Rothbarth approaches supported by the empirical translation of the theory through equations in Section 4. Next, a description of the data used will be explained in Section 5, followed by a section of regression results (Section 6) and child share cost results (Section 7); and finally, we will conclude the paper in Section 8.

2. LITERATURE REVIEW

In this section an introduction to the current estimates that are found in the literature will be presented, followed by a general criticism of the models that will be later considered. This section is important in that it will hopefully give the reader a sense of where the findings in this particular thesis work may (or may not) fit in with other relevant investigations.

2.1 Current Estimates of Total Expenditures on Children

Espenshade (1972) made a strong impact on the literature pertaining to the estimation of the cost of children with his book, *The Cost of Children in Urban United States*. A rather detailed literature review of pertinent contributions up through 1972 is contained in the book, as well as the foundation of the Henderson-based methodology for his estimation results that were to be used in said book. Later, Espenshade published the more updated version of this earlier study in his 1984 contribution, *Investing in Children*. The methodology found in the 1984 contribution is essentially the same found under the Engel Model in this paper – using Engel's presumption that food at home can stand in for a measurable standard of living variable, Espenshade first allows living standards to vary over the life cycle and second, he dismisses the idea that expenditures on children do not vary with birth order. These two assumptions are then incorporated into a model that will be used to estimate how much a household with children needs to be compensated to attain pre-children welfare levels. The main results from Espenshade (1984) suggest that expenditures on children vary according to household income, the education level of the parents, the employment status of the mother, and family size. Additionally, he finds that as children age, they require an increase in expenditures, but there is a reduction in expenditures per child as more children are added to the household. Espenshade further provides a rather well developed list of reasons for choosing food at home as the iso-prop measure[4] and rejecting expenditures on adult clothing in its place; among them, the expenditure should feature prominently in total household expenditures in order to avoid the introduction of random measurement error, and adult clothing (in his sample from the 1972-73 CEX) only accounts for 8-9 percent. Finally, after running several regression models, Espenshade asserts that the best regression should coincide with the following criteria: (i) all of the signs on the age-sex coefficients should be equal (not necessarily positive or negative) and opposite the sign on total expenditures and (ii) the age-sex coefficients and the total expenditures coefficient should all be statistically significant. Actual child share estimation results from this particular study can be found in Table 2.1.

While Espenshade can claim a virtual monopoly over the competition of initial entrants into the market of estimation, the work of Dr. David Betson of the University of Notre Dame is now gaining some ground. An initial report conducted for the U.S. Department of Health and Human Services (1990) introduced Betson to the child support community, wherein he estimated expenditures on children using 1980-86 CEX data and the employment of five different well-known models: the per capita method, the Engel method, the iso-prop method, the Rothbarth method, and the Barten-Gorman (see page 11 for a description) model. Results from the Engel method differed little from those found in Espenshade (1984), and these numbers were later updated in an unpublished document from 2000 that used data from the 1996-98 CEX (results can be found in Table 2.1). As for the Rothbarth method, Betson is an avid proponent of this method over the others; recent estimations (Betson, 2006) can be found in Table 2.1. General conclusions from Betson's efforts support the other findings that total expenditures on children rise as the number of children rise, but the average cost per child falls, and older children tend to require more resources.

The U.S. Department of Agriculture (USDA) has also provided estimates on expenditures on children since 1960 (Lino, 2007). The estimation techniques used in their studies are also based on multivariate regressions; however, these studies use the per

[4] Bassi and Barnow (1993) provide the following definition of iso-prop: short for iso-proportion, these estimators are rooted in the assumptions that "(1) the percentage of family expenditures devoted to a particular type of consumption (e.g., food…) serves as a proxy for the family's utility, and (2) the family's expenditure patterns are separable." (p.481)

capita method[5] of estimation (as opposed to the marginal estimation used in Espenshade) and justify this use via the following: Lino asserts that the marginal approach is not supported by empirical findings and is merely based on theory. Indeed, the basis of his argument is that marginal techniques depend on the development of equivalency scales (the key assumption that families who spend proportionately equal amounts on food share equal welfare levels has yet to be proven[6]). Furthermore, Lino also claims that many of the marginal approaches also fail to take into account substitution effects. To show that an effort was made to embrace the marginal approach, in a 1995 study by Lino and Johnson, which Lino cites in the USDA (2007) report, estimation of expenditures on children were conducted by using housing, transportation, and miscellaneous expenditures as the welfare equivalency measure under a marginal cost approach. According to Lino, the Rothbarth technique resulted in estimates that were "unrealistically low" (2007, p.12) while the Engel technique yielded estimates that were below the USDA findings.

Using data from the 1990-92 CEX and inflating to 2007 dollars through the CPI, the USDA report utilized the 1994 food plans published by the USDA to allocate food expenses among household members, and finds that housing accounts for the largest share of total expenditures on children and (as with Espenshade, 1984) expenditures on children rise as the children's ages rise. Numerical results illustrating this behavior can be found in Table 2.1.

For a response to Lino's concerns regarding the use of per capita methods over marginal estimation techniques one should begin with Lazear and Michael (1988), whose entire book, *Allocation of Income within the Household*, contains interesting evidence that questions the logic behind the assumption that household members consume equal amounts. While the assumption itself is not necessarily at the heart of the matter, it is the effects of utilizing such an assumption in an equation estimating expenditures on children

[5] The per capita approach does not attempt to attach unique expenditures per household member. Instead, it essentially considers averages in its calculations, by taking total household expenditures and dividing it equally amongst household members. Thus, adults and children are assumed to consume, on average, equal amounts. The marginal approach attempts to attribute incurred additional household costs to the additional member of the household.

[6] Additionally, Nelson (1993) would append that such a narrow focus on household income potentially ignores other components that tend to enrich a family's standard of living, such as quantity and quality of publicly provided goods (schools, roads, public transportation, etc.).

that can lead to misinterpretations. Using data from the 1972-73 CEX Lazear and Michael find that children do not command the same expenditures as adults within a given household; indeed, they go on to state that the per capita method greatly understates the amount attributed to an adult, since their results suggest that an adult receives 2.5 times the amount dedicated to a child. Moreover, Lazear and Michael (1988) conclude that on average, a child receives about 40 percent as much of the household income as an adult in the same house would. The authors do stress that allocation patterns vary considerably with the characteristics of the household, but do find some evidence that suggests in general, girls tend to cost more than boys. It should also be noted that the basic model employed during estimation utilized adult goods (i.e., clothing, tobacco, and alcohol) as the stand in variable for welfare levels, and the authors do agree with Espenshade on the limitations such a relatively small expenditure relative to total expenditures can present.

Table 2.1 Comparison of Results from Engel, Rothbarth, and Per Capita Approaches

Author and Year	Model	Estimated Child Expenditures as a Percent of Total Household Expenditures		
		One Child	Two Children	Three Children
Espenshade (1984)	Engel	24%	41%	51%
Betson (2000)	Engel	30%	44%	52%
NATSEM (2005)	(enhanced) Engel	11-17%	18-32%	24-34%
Lazear and Michael (1988)	Rothbarth	11%	19%	25%
Betson (2006)	Rothbarth	25%	37%	44%
USDA (2007)	Per Capita	24%	37%	56%

In a similar study, undertaken by Percival and Harding (2005) for the National Centre for Social and Economic Modelling (NATSEM), the authors attempt to estimate expenditures on children using what is the equivalent to the U.S. CEX survey in Australia, the Household Expenditure Survey (HES), from 1998-99. What is perhaps noteworthy is that the proxy for welfare levels is a combination of the following basket of goods: food at home, fuel and power expenses, household non-durables at home, postal, telephone, and telegram costs, and personal care products (all deemed necessities).

Notable missing items include clothing, housing, and health. The authors justify these omissions by stating that (i) the inclusion of housing costs would create distortions (e.g., housing costs diminish over a life cycle), (ii) clothing is considered a luxury item (and thus, harder to separate clothing that qualifies as necessary vis à vis clothing purchases that are a luxury), and (iii) Medicare in Australia covers most households' health expenses. Results indicated that expenditures varied with the age of the child, the income level of the parents, and the number of kids in the household, while the trend in expenditures was also similar to that found by the studies mentioned earlier: younger children require less resources and average expenditures per child tend to decrease with the addition of children. Numerical results supporting this behavior can be found in Table 2.1.

2.2 Criticisms

Because it is obvious that estimation of expenditures on children not only depends heavily on the defining model but also on the data set used, it is important to bring forth some criticisms that have been voiced in the academic literature surrounding these approaches.

Deaton and Muellbauer (1986) suggest that the Engel method cannot possibly be correct in its assumptions that food shares are a valid stand in for the household's standard of living, even if the two most common assumptions[7] associated with the food share model are accepted. Moreover, according to these authors, the Engel method will most likely overstate expenditures on children, citing results that suggested one child would require 82 percent of the resources required by an adult (recall, Lazear and Michael (1988) suggested this number was closer to 40 percent, on average), while two kids would each require 77 percent of the resources required by an adult. As for the Rothbarth approach, Deaton and Muellbauer argue that results would most likely understate the 'true' cost of children. Citing a study by Cramer (1969), the authors point out that tobacco and alcohol do not appear to be very responsive to changes in income,

[7] The two commonly accepted assumptions for the Engel approach include the following: an increase in family size is associated with an increase in the budget share attributed to food and secondly, an increase in total expenditures is followed by a decrease in the budget share devoted to food.

which would, in turn, make it difficult to establish the income effects required for equivalency scales. Yet another primary argument Deaton and Muellbauer present against the Rothbarth approach is that substitution toward adult goods in a household with children will not be captured, leading to estimates that are too small. These authors supplement their criticisms with a potential solution: the Barten-Gorman method[8], which could be considered a generalization of both the Engel and Rothbarth methods. However, they do note that in practical applications, the parameters required for the Barten-Gorman model are extremely difficult to estimate, which inevitably leads to the current popularity of the Engel and Rothbarth methods.

Furthermore, Bassi and Barnow (1993) would also support these conclusions, stating that while a bias is expected in either the Engel or Rothbarth cases, the direction of the bias is believed to be known. Indeed, yet another critical assumption in the Engel Model that fails to be empirically measurable and potentially contributes to the bias is the Engel condition of separability. Bassi and Barnow (1993) explain that the foundation of this assumption lies in accepting the equivalence of the child-attributed percentage of total household expenditures on *nonfood* items with the child-attributed percentage of total household expenditures on *food* items. However, intuitive reasoning may deem this assumption incorrect – children are perhaps, as Bassi and Barnow put it, "relatively 'food-intensive'" (1993, p.483). This would coincide with the belief that the Engel approach overestimates expenditures on children. The authors conclude, then, that the Engel estimator should be considered an upper bound while the Rothbarth estimator, while perhaps too low, should be adjusted upwards to reflect more realistic numbers. Nelson (1992), however, asserts that empirical evidence suggests economies of scale are indeed present within the household, do affect consumption patterns, and thus, can result in relative price changes, which, when not accounted for (as in the Rothbarth model), leads to *over*estimation of the compensation needed to restore the household to pre-children welfare levels. The author does add, however, that the direction of the bias cannot be known for sure, since "the degree to which adult goods may be complementary

[8] The Barten-Gorman method, as it is now called, is an updated version of the model presented by Barten (1964), which is a utility-based approach wherein the proxy for household standard of living results from weighted commodity purchases. In this type of model, income and substitution effects on household consumption as caused by the presence of children are allowed; however, the model inherently can only account for parental utility, not that of children, as suggested by Nelson (1986).

to goods with high economies of scale" is not known (1992, p.301). The essence of both studies, therefore, is to bring forth the issue of separability inherent in either model, and the potential invalidations of assumptions introduced via the existence of implicit price substitution effects.

A third study that takes this point of contention even further shows that empirical evidence actually contradicts theoretical assumptions. Deaton and Paxson (1998) assert that, in general, if household scale economies were to exist, then theoretically speaking, households should be better off as family size rises, holding per capita resources constant. This, in turn, would imply that if substitutions towards cheaper household public goods fail to take place, then the income effect would lead larger households to have higher per capita consumption of private goods. In fact, the Deaton and Paxson (1998) find the exact opposite during their empirical investigation. Holding per capita total expenditures of the household constant, they find that per capita food expenditures fall as family size increases. This being the case, Engel's assumptions, which encapsulate positive estimates of economies of scale, could indeed be incorrect. Accordingly, this is the very paper Betson cites in his work as compelling evidence against the Engel approach, and, consequently, in his opinion, is what renders the model by Rothbarth more plausible (Betson, 2000).

2.3 Concluding Remarks

This section has shown that there is a tremendous amount of disagreement, both theoretically and empirically, in the quest to address the estimation of child costs. While this study does not attempt to explicitly investigate in particular any one of the concerns mentioned above, these contributions to the literature do assist in being able to speak more confidently on the importance and relevance of the findings in this thesis. The following section will introduce the theoretical models that will be used in the remainder of this study.

3. THEORETICAL MODEL

As mentioned previously in this study, there is a great deal of difficulty extracting specific consumption patterns for the individual members of a household. Thus, it is common practice to instead consider the household as one unit and compare it with similarly well-off households of a different internal composition, in order to investigate the possibility of attributing particular costs to certain members of a household. However, a new obstacle is introduced given this approach: how does one chose a variable that can best attest for standard of living? The economic notion of utility is immeasurable (and not necessarily valid when the unit under investigation is not a sole consumer but a household) and income itself does not speak accurately of the concept of standard of living (one needs only to visualize just how far a yearly salary of $25,000 can go for a family of one versus a family of four). Indeed, a way to get around this obstacle is by directly measuring a variable that can stand in as the indicator for the household's standard of living, and use the differences in expenditure patterns among similarly well-off households to explain which family member may be the source of certain costs.

In our case, we are searching specifically for those costs that may be attributable to the children of the household. Thus, an estimate of expenditures on children can be surmised, and a resulting cost of children estimate can be proposed. It is now a matter of which variable one should choose to most accurately measure standard of living; justification for such a choice should be grounded in the reasoning that this variable should be able to withstand different sets of sample data over time with little change in its ability to perform within an equation. The similarities between the foundations of the Engel and Rothbarth Models were presented in the Section 1 of this paper; accordingly, the choice of which variable should in fact be the proxy for standard of living is the point of diversion for these models. The objective of this chapter, thus, is to present the theoretical underpinnings behind two of the most commonly used methods for estimating expenditures on children, the Engel and Rothbarth Models, including a quantitative and qualitative sketch of both models.

3.1 Engel Approach

Ernst Engel (1857), in his search for a variable that could withstand the changes time might bring to changes in household consumption patterns, understood that a solid proxy for standard of living could most likely be found in an item that is a necessity for every household and consequently, could, in fact, be found in every household: food. Engel came upon the discovery that expenditures on food for at home consumption seemed to follow a well-ordered pattern, wherein the rise of total expenditures is be accompanied by a consequent fall in the percent of expenditures reserved for food. Secondly, he also noted that an increase in the size of the family should also be seen in conjunction with an increase in the percent of expenditures dedicated to food. To marry these theoretical notions within this study, the final observation to note from Engel's contributions is that from these assumptions, one can conclude that expenditures on food appears to be a strong candidate for the well-being of a household – that is, a family is assumed to be equally well-off if they exhibit equal amounts of expenditures on food for at home consumption.

In an approach guided by the 1990 contribution of Betson, the following quantitative sketch will assist in illustrating the mathematical relationships among the important variables under consideration. The ideas presented by Engel can be captured in the following equation: Let Ω be a function of X, total expenditures, and N, the number of children in a household, such that the share of total expenditures devoted to food at home, Y, is then computed by

(1) $$\Omega (X, N) = Y.$$

To capture the estimate of expenditures on children, we introduce a new variable, C, that denotes the expenditures devoted to the presence of a child in the household. By incorporating this variable into the equation, the Engel approach says that the cost of one child is determined from the following modified equation:

(2) $$\Omega (X, N_1) = \Omega (X - C, N_0)$$

where N_1 refers to a household with one child and N_0 refers to a household with no children.

A qualitative graphical representation of the relationship described by equation (2) can be found in Figure 1. In this figure, the vertical axis (Y) indicates the share of total expenditures dedicated to food at home, while the independent variable, total expenditures (X), is depicted on the horizontal axis. As shown in this sketch, a family consisting of a household with no children is denoted with the variables containing the subscript zero (0), while a household containing one child is denoted with the variables containing a subscript of one (1). Based on the relationship presented in equation (2), by comparing a family that has equivalent expenditures on the share of food at home (in other words, enjoying the same standard of living), on the graph denoted $Y_0 = Y_1$, the difference between total expenditures, found by subtracting X_0 from X_1, would represent the cost of one child, denoted C.

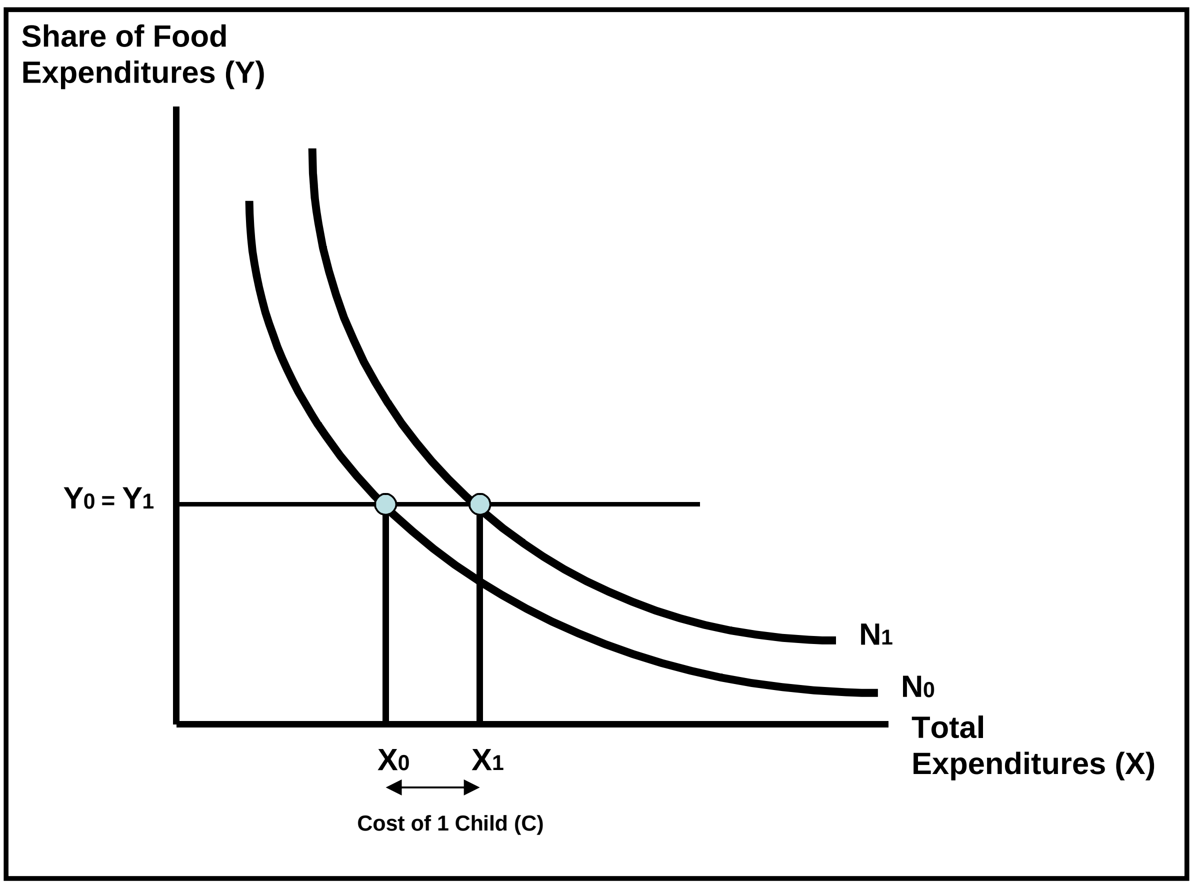

Figure 1: Graphical Representation of Engel Approach

3.2 Rothbarth Approach

A second approach that relies on a similar structure as the Engel method is the model proposed by Erwin Rothbarth (1943), which is also a very commonly used method of estimating expenditures on children. Rothbarth postulated that after the basic necessary needs are met, households have a given level of excess income that could be distributed among a variety of goods, including savings accounts, entertainment options, as well as a host of other luxury items. Most current studies, however, have consolidated this list into a mere handful of 'adult items' consisting of alcohol, tobacco, and adult clothing. The proxy for standard of living, in the case of the Rothbarth Model, is the total expenditures on these adult items; more specifically, in this study, as guided by Betson's 2006 Rothbarth study, the potential list of adult items will be restricted further to include only those expenditures on adult clothing. This is indeed the point of diversion between the Engel and Rothbarth Models. As will be shortly discussed, some of the expected relationships will change, as a result.

The Rothbarth approach should produce results that will graphically look slightly different than those produced using the Engel method (as presented in Figure 1). That is, for this model to be valid, it is first assumed that the relationship between the shares of expenditures dedicated to adult items and total expenditures is of a positive nature (rather than negative, as in the Engel model). Thus, as total expenditures rise (or as the household's standard of living improves), a consequent rise in the share of expenditures reserved for adult goods (a luxury item) should be observed. Secondly, as the household size increases, one should see a decrease in the amount of expenditures related to adult items (a negative relationship, as opposed to the positive relationship observed in the Engel Model).

A mathematical framework is also appropriate here for illustrative purposes and is quite similar to the Betson-guided presentation of the Engel Model found in Section 3.1. In this case, let Ψ be a function of X, total expenditures, and N, the number of children in a household, such that the share of total expenditures devoted to adult items, A, is computed by

(3) $\Psi(X, N) = A$.

To capture the estimate of expenditures on children, we introduce a new variable, C_1, that denotes the expenditures devoted to the presence of a child in the household. Incorporating this variable into equation (3), the Rothbarth approach says that the cost of one child is determined from the following modified equation:

(4) $$\Psi (X, N_1) = \Psi (X - C_1, N_0)$$

where N_1 refers to a household with one child and N_0 refers to a household with no children. The similarity to the Engel model should be easily noted. It is important to remember that it is the relationships between variables that change.

A qualitative graphical representation of the relationship described in equation (4) can be found in Figure 2. In this figure, the vertical axis denotes the share of total expenditures devoted to adult items (A), while the independent variable, total expenditures (X), is depicted on the horizontal axis. As shown in the sketch below, a family consisting of a household with no children is denoted with the variables containing the subscript zero (0), while a household containing one child is denoted with the variables containing a subscript of one (1). Based on the relationship previously discussed, by comparing a family that has equivalent expenditures on the share of adult items (in other words, enjoying the same standard of living), on the graph denoted $A_0 = A_1$, the difference between total expenditures, found by subtracting X_0 from X_1, would represent the cost of one child, denoted C_1.

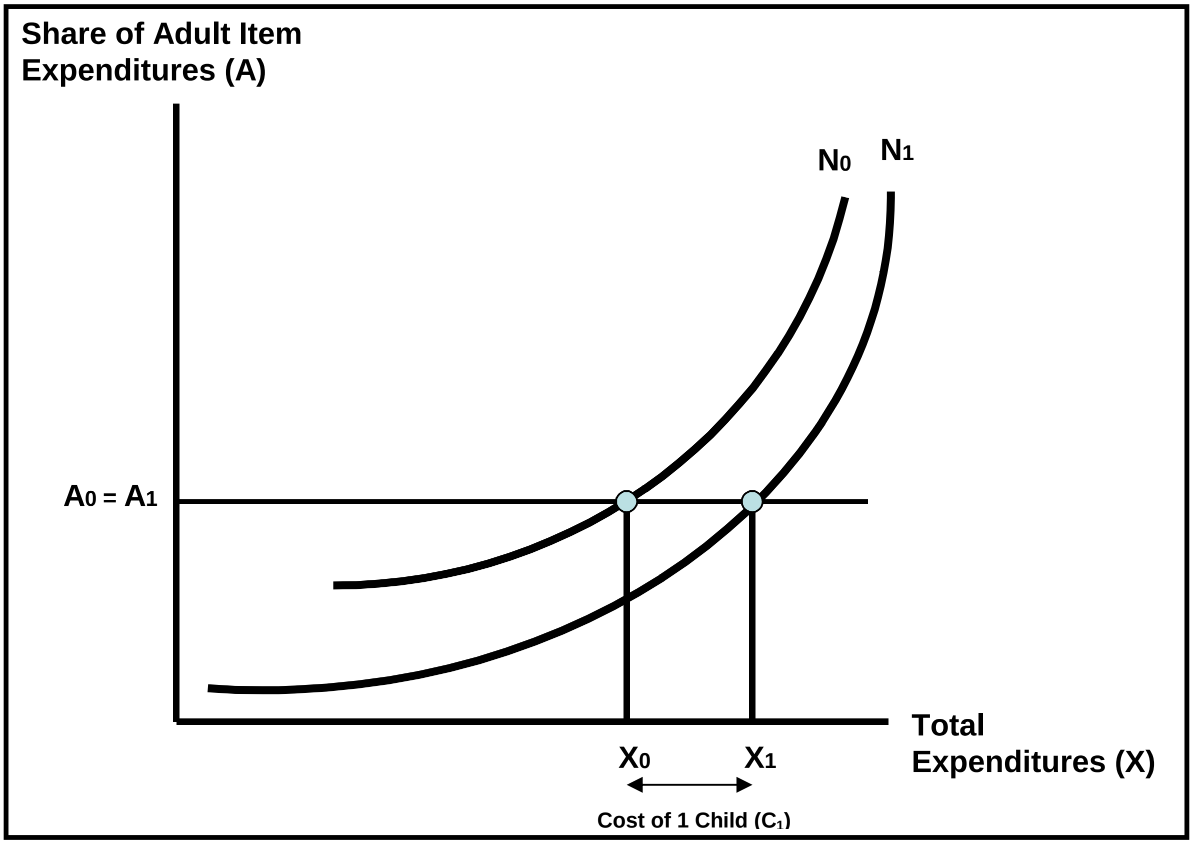

Figure 2: Graphical Representation of Rothbarth Approach.

3.4 Concluding Remarks

We have now introduced the theoretical foundations behind either model, as well as presented some of the primary differences in assumptions that will dictate how the regression results might vary. The following section will take these theoretical musings and translate them into the actual regression equations that will be used to obtain child cost estimation results, as well as the results for testing model robustness.

4. EMPIRICAL MODEL

While the main motivation behind the previous section was to introduce the relationship between the variables to be studied, the principal purpose behind this section is discussing the translation of the quantitative sketches presented earlier into an actual regression that can be tested. Thus, in the subsequent section we will first look at the regression equation for the Engel Models, and next, the regression equation for the Rothbarth Model.

4.1 Engel Equation

In this sub-section we are going to build on the basic equation, (1), found in Section 3.1, and expand some of the independent variables that will assist in determining the dependent variable, the share of food at home. While this translation is not unique, recall that one of the primary goals of this thesis is to replicate the work presented by Betson; therefore, this regression equation is essentially the same one detailed in his studies.

For the Engel approach, the value of total expenditures (the variable X from equation (1)) is substituted into the following equation along with a collection of what are considered other relevant variables in order to predict the per capita food share of total expenditures:

$$\text{(5)} \qquad \ln\left(\frac{F}{(1-F)}\right) = \theta \ln(N) + \delta \ln(\frac{S}{N}) + \beta \ln(\frac{S}{N})^2 + \alpha(K) + \gamma(X)$$

In the above equation, the dependent variable, ln (F/(1-F)), is defined to be the natural log of the ratio of total food-at-home expenditures, F, to one minus total food-at-home expenditures, 1-F. The corresponding food variable from the CEX data source is the variable that is used to measure expenditures on food-at-home consumption. The share of per capita food-at-home expenditures is assumed to be a linear function of the following: (i) the natural log of family size, $\theta\ln(N)$; (ii) the natural log of per capita total expenditures (total expenditures divided by family size), $\delta\ln(S/N)$; (iii) the square of this term, $\beta\ln(S/N)^2$; (iv) a collection of variables that gives the breakdown of the proportion of kids in a family based on the child(ren)'s age(s), $\alpha(K)$; and finally, (v) a collection of demographic variables specific to the adults found in the family, $\gamma(X)$. The inclusion of the exponential term, $\beta\ln(S/N)^2$, is necessary to allow for a nonlinear relationship between food-at-home expenditures and total expenditures.

Recalling the discussion found in Section 3.1 regarding Engel's assumptions, for equation (5) to hold, the coefficient δ on the per capita total expenditures variable should be negative. That is, as total expenditures rise, the share of expenditures dedicated to food at home should fall. Additionally, the collection of variables pertaining to the proportion of kids in the household, $\alpha(K)$, should all be positive, since it is assumed that, all other things constant, as the number of kids within a household rise, expenditures on food at home should consequently rise. This same line of reasoning can be applied to the coefficient on family size, θ, since an increase in the number of individuals in a household is effectively the same as adding children. Thus, θ should also be positive.

4.2 Age-Invariant Equation

Equation (5) in Section 4.1 serves as the fundamental basis for the regression equation developed for what we are calling the "Age-Invariant" Model. This interpretation of the Engel approach was first introduced by the Department of Economics at Florida State University in 2004 in a report to the Florida Legislature on updating Florida's child support guidelines. We are again employing it in this thesis, in effort to try and investigate further the reliability of the Engel approach. By simplifying some of the independent variables, we hope to isolate the strongest contributors to the calculation of the dependent variable, and gain a better understanding of how the Engel approach responds to the data.

The major differences between the two versions of the Engel approach include the following: (i) the collection of child-age proportion variables, $\alpha(K)$, is removed from equation (5); in its place, a single collapsed variable to account for the presence of an individual child (up to five possible kids in a household), regardless of age, is included and (ii) the term for family size, θln(N), is also removed; including this term would result in collinearity between the individual child variables introduced in (i) and the variable measuring family size. This slight modification does not otherwise change the sign expectations on any of the other variables, including the child variables (which are still assumed to be positive).

4.3 Rothbarth Equation

For the Rothbarth approach, we will return to the guidance of Betson's earlier contributions; thus, the similarities with equation (5) will surely become apparent. In this case, once again, the value of total expenditures (the variable X from equation (3)) is substituted into the following equation, along with a collection of other assisting independent variables, in order to predict total expenditures on adult goods (specifically, adult clothing):

$$(6) \qquad \ln(A) = \theta \ln(N) + \delta \ln(\frac{S}{N}) + \beta \ln(\frac{S}{N})^2 + \alpha(K) + \gamma(X)$$

Readers will recognize the single difference between equation (5) and equation (6) as the dependent variable, ln (A), which, in equation (6) is defined to be the natural log of the *total* amount of expenditures devoted to adult clothing. The corresponding adult-items variable from the CEX data source had to be constructed by totaling expenditures from the following variables, and creating a single adult clothing expenditure variable: expenditures on women's clothing (females aged 16 and up) and expenditures on men's clothing (males aged 16 and up). The remaining variables in equation (6) are as described in Section 4.1 of this study.

The Rothbarth method, based on the assumptions noted in Section 3.2 of this study, should produce the *opposite* effect on the coefficient on total expenditures, δ, as the Engel Model; indeed, if the assumption that adult clothing is a luxury good is correct, then we would expect to see an increase in total expenditures on adult clothing as total expenditures rise. This coefficient, then, should be positive. Where these two equations (eq. (5) and (6)) should also differ is in the coefficient estimates on the number of kids in a family, $\alpha(K)$. Given that we are estimating total expenditures on adult items only, all other things constant, when kids are introduced into the household, the portion of expenditures dedicated to adult items should fall. Therefore, α should be negative, on every category of children's ages, with the exception, perhaps, being the kids aged 16-17, who are technically considered adults in the CEX data source. Furthermore, because of the imposition of the data restriction that only those households consisting of two married adults were included in the final sample (see section 5.2 for further details), the coefficient on family size, θ, should also be negative (different than the hypothesized results for the Engel equation). That is, as the size of the family increases, due to the introduction of children into the household, all other things constant, total expenditures dedicated to adult items should fall.

The remaining collection of demographic variables specific to the adults in the household, $\gamma(X)$, is included in all three equations in order to account for differences in the makeup of the family. These include a variety of characteristics, such as race, education, and labor force decisions. While these variables may help to explain some of the differences in estimating total share of expenditures spent on either food-at-home or adult items, it is ultimately assumed that these variables do not vary with the number of children in the household, and should consequently not have an impact on estimating expenditures on children in either the Engel or Rothbarth cases.

4.4 Definition of Variables Used

In order to implement the equations denoted in this section, the definition of variables included in the respective regressions can be found in Table 4.4.

Table 4.4 Variable Definitions	
Variable	**Definition**
LNFSIZE	Log of family size
LNPCTEXP	Log of per capita total expenditures
LNPCTEXP2	Square term of log of per capita total expenditures
CK02	Number of children 0 to 2 years old divided by family size
CK35	Number of children 3 to 5 years old divided by family size
CK612	Number of children 6 to 12 years old divided by family size
CK1315	Number of children 13 to 14 years old divided by family size
CK1617	Number of children 15 to 17 years old divided by family size
CA1825	Number of adults 18 to 25 years old divided by family size
CA2635	Number of adults 26 to 35 years old divided by family size
CA3645	Number of adults 36 to 45 years old divided by family size (reference group, omitted in regression)
CA4655	Number of adults 46 to 55 years old divided by family size
CA5660	Number of adults 56 to 60 years old divided by family size
BLACK	1 if the Head was black, 0 otherwise
HD_NO_HS	1 if Head's education was less than 12 years, 0 otherwise
HD_COLL	1 if Head's education was greater than 12 years, 0 otherwise
SP_NO_HS	1 if spouse's education was less than 12 years, 0 otherwise
SP_COLL	1 if spouse's education was greater than 12 years, 0 otherwise
W_WORK	Weeks worked by spouse divided by 52
FTIME	1 if the spouse worked more than 30 hours per week, 0 otherwise
TWOERN	1 if both adults worked, 0 otherwise

4.5 Concluding Remarks

We have reached the point in this study where we are now ready to introduce a description of the data sets used in the calculation of the empirical equations presented in this section. The following section will give a detailed explanation of the restrictions deemed necessary in cleaning up the data obtained from the CEX.

5. DATA

The importance of this section cannot be overstated. That is, the central purpose of this thesis is to prove robustness in the Engel and Rothbarth Models; our plan for achieving just that is by testing two similarly constructed data sets that differ only on one condition. Thus, in this section we detail and justify the initial restrictions made to the untouched data source first obtained from the Bureau of Labor Statistics (BLS), and next, we introduce the restriction that will define the two differing data sets to be used in running the regressions. We will conclude with some general descriptive statistics that will help illustrate some of the differences among the two data sets.

5.1 Consumer Expenditure Survey (CEX)

The source of the data used in this study is the Consumer Expenditure Survey (CEX), which is conducted by the U.S. Census Bureau and published by the U.S. Bureau of Labor Statistics (BLS). This is essentially the only survey that captures a national sample of U.S. families. As such, the CEX attempts to collect data on a wide variety of household characteristics, including family spending patterns, household income, and demographic attributions of U.S. families through the annual survey method. Furthermore, the survey itself consists of two parts: (a) an Interview Survey, which is conducted quarterly, and focuses on monthly out-of-pocket expenses (for example, expenditures pertaining to housing, apparel, transportation, health care, and entertainment) and (b) a Diary Survey, conducted weekly, and focuses more on the expenditures of items purchased frequently (for example, food, tobacco, personal care products, etc.).[9] This study only uses the data from the Interview Survey.

[9] CEX Overview, http://www.bls.gov/cex/csxovr.htm

For the Interview Survey, every consumer unit[10] (roughly around 7,000 per CEX) is interviewed every three months in the course of a little over a year, totaling to five calendar quarters. The data used in this study was collected from the first quarter of 1999 to the last quarter of 2001.[11] A hypothetical annual data set was constructed, where each household was given a unique ID and then linked across all quarters. The only variable not to be measured quarterly was income, and thus, had to be constructed as the arithmetic average of the quarterly data. Additionally, the number of children in a household was also averaged across quarters, and as a result, some households may indeed have fractional children, if a child was present in the household for less than the full year.

While interviews are conducted quarterly, there may be missing pieces of information due to the failure of a response from a consumer unit, particularly, for example, the failure to state expenditures on food for a certain quarter, which can thus render a household with only partial data. It is this point of divergence that we are going to use as a distinguishing restriction between our two data sets. There is no strong economic rationale to exclude from the sample those households that responded to less than 3 of the 4 interview attempts unless the households that contain missing interview responses are somehow correlated with the dependent variable featured in either model; this is not assumed to be the case. As a result, we are directly testing the robustness (as described in Section 1 of this study) of all three models introduced in Section 4 by subjecting each to two data sets that differ only by this very component. After all relevant deletions were incorporated, the final sample was divided into (i) a "complete" data set that included all the final households regardless of response rate and (ii) a "partial" data set that included only those final households that reported at least three responses. This is

[10] Defined to be one of the following: "(1) All members of a particular household who are related by blood, marriage, adoption, or other legal arrangements; (2) a person living alone or sharing a household with others or living as a roomer in a private home or lodging house or in permanent living quarters in a hotel or motel, but who is financially independent; or (3) two or more persons living together who use their incomes to make joint expenditure decisions." CEX Frequently Asked Questions, http://www.bls.gov/cex/csxfaqs.htm#q1.

[11] This is about half of the sample that is used in the Betson (2006) study; that is, there are three years worth of data here, while Betson uses roughly six years worth of data (1998 through the first quarter of 2004) (Betson, 2006, p. 8).

the major difference with the Betson (2006) study, wherein the only sample he used was a "partial" data set, deleting all observations that had less than three responses.

5.2 Explanation of Deletions

Beginning with a sample size of 34,893 households (91,379 observations), a handful of deletions was necessary in order to arrive at an appropriate sample; Table 5.2 summarizes the extent of all deletions. The criteria for this process were guided by the noted implementation of deletions in the Betson (2006) study and are based on the following:

(i) only those households that consisted of a husband and wife, between the ages of 18 and 60, were included

(ii) only those households that contained no other adults (persons over the age of 18) present were included, even if these adults were the children of the parental units

After solely applying the above two restrictions, a 27 percent reduction from the initial sample size occurred, and only 9,461 households remained.

No mention was made in the Betson (2006) report as to the exclusion of households that have zero, negative, or missing income. Thus, no deletions were made based on this restriction. On the other hand, the BLS does modify some of the data in ways that may unintentionally bias results. For example, in order to avoid revealing survey participants due to sensitive responses, some of the data points (in this instance we are solely focusing on the data points having to do with after-tax and before-tax income) are topcoded.[12] Any observations testing positive for this criterion were consequently deleted; in this instance, an additional 756 deletions were noted. While Betson (2006) does not explicitly state that topcoded variables were deleted during this data cleansing process, he does recommend a cautious interpretation of results when total expenditures exceed $75,000.

[12] Topcoding refers to the practice of identifying those data points that fall outside recommended critical values and replacing them with a so-called topcoded value (e.g., the value of the mean of all outlying observations) (CEX 2001 Interview Survey Public Use Microdata Documentation).

The definition of total expenditures also had to be tweaked in order to coincide with theoretical assumptions. The research undertaken in this study is specifically looking at the consumption patterns of households, and any expenditure that is not directly associated with consumption by a household member must thusly be excluded. The BLS includes in its measure of expenditures cash contributions that are made to members outside the household. This item, therefore, had to be excluded. Additionally, the BLS includes contributions made to Social Security and pension plans. Because this technically constitutes savings and not consumption, this had to be excluded as well. Total expenditures, then, as defined above, was next averaged across quarters and multiplied by four to construct an estimate of annual expenditures. It is presumed that this is the same construction technique utilized by Betson in his studies.

A summary of average total expenditures by household composition and data sets can be found in Table 5.3.2. It is perhaps important to note that average total expenditures for either of the data set used in this study are consistently lower than the averages found in the Betson (2006) study. This could be due to a variety of reasons; chief among them the three years' worth of data we are excluding, but also, there may be some other expenditure exclusion that Betson may have inadvertently omitted from the (2006) list.

Table 5.2: Number of Sample Deletions by Reason	
Total Number of Households in the 1999-2001 Consumer Expenditure Survey	**34, 893**
Reduction for:	
a) Non-husband and wife households	-22,126
b) Spouse # 1 not between 18 and 60 yrs old	- 3,099
c) Spouse # 2 not between 18 and 60 yrs old	- 239
d) Households with topcoded income variables	- 756
e) Households with 6 or more children	- 28
f) Households with zero expenditures on clothing	- 1,386
Usable "Complete" Sample	**7, 259**
g) Households with less than 3 completed interviews	- 3, 078
Usable "Partial" Sample	**4, 181**

Along similar lines of reasoning behind the removal of topcoded variables, it was decided that households that contained six or more children would also bias results, and thus had to be deleted. 28 households were deleted due to this reason. And finally, the last restriction noted in the table, according to Betson, that had to be implemented was the exclusion of households that had zero expenditures on adult items: 1,390 additional observations were dropped.

5.3 Descriptive Statistics

Following the completed sequence of deletions as noted in Section 5.2, a set of descriptive statistics is useful in comparing all sample data sets. In terms of distribution of family sizes, (i) both data sets in this study contain the greatest number of families in the “0” children household composition group, and (ii) of the households that do contain at least one child, both data sets have a greater number of households that contain two children relative to those households that contain either one or three or more children. A summary of this information can be found in Table 5.3.

Table 5.3: Distribution of Observations by Number of Children per Family

	Number of Kids				
Sample Data Set	**0**	**1**	**2**	**3**	**Total**
Complete Sample	2,714	1,542	1,960	1,043	7,259
Partial Sample	1,504	874	1,172	631	4,181
Betson (2006) Sample	3,338	1,778	2,611	1,518	9,245

Indeed, a more useful picture may be found in the percent of households committed to each category under ‘number of kids’. For example, despite the reduction in sample size from the “complete” to the “partial” used in this study, the resulting sample data sets are relatively analogous in percentage distribution of family composition. Similarly, the numbers indicate that the sample used in Betson (2006) also adhere to this allocation. Table 5.3.1 contains a summary of this information.

Table 5.3.1: Percent Distribution of Observations by Family Composition				
	Number of Kids			
Sample Data Set	**0**	**1**	**2**	**3**
Complete Sample	37%	21%	27%	14%
Partial Sample	36%	21%	28%	15%
Betson (2006) Sample	36%	19%	28%	16%

Total expenditures as calculated by the BLS include a subset of other various expenditures; among them, expenditures on food, housing, transportation, etc. The most relevant variables to this study include food at home and clothing, as they are being considered the proxies for household welfare within their respective models. Thus, a general statistical description of how each changes according to number of children in the household is useful for gaining a better understanding of trends. Table 5.3.2 contains a listing of the share of average expenditures devoted to a smaller subset of total expenditure categories differentiated by household size.

In general, average total expenditures increase as household size increases; this is similar to findings in other studies and is supportive of Engel's assumptions. The trends for the variables known as "adult goods" in either sample (tobacco and alcohol) show a decrease as the number of children present within a household rises, as to be expected. Adult clothing exhibits a similar trend; however, it is important to emphasize at this point that this is indeed the variable being employed in the Rothbarth model, and yet it does not quite make the three percent mark for average share of total expenditures, even at its maximum. The variable being called upon in the Engel model, food at home, does, instead, reach a high of almost 14 percent of total expenditures on average. This perhaps suggests a bit more resistance to random sampling error, as unfortunately might not be the case with clothing, due to the small numbers affecting its average share. Furthermore, food at home tends to increase with number of children in a family, as to be expected. There are some slight differences in percentages when comparing the "partial" and the "complete" data set for the specific proxy variables. Firstly, the average total expenditures found in the partial data set tend to be somewhat smaller than the average total expenditures found in the complete data set. Secondly, adult clothing is a consistently smaller share of total expenditures in the partial data set as compared to the

Table 5.3.2: Selection of Average Spending Components by Family Composition				
	Number of Kids			
	0	**1**	**2**	**3**
Average Total Expenditures				
Complete Sample	$38,179	$41,176	$45,314	$45,942
Partial Sample	$36,597	$40,281	$44,521	$44,199
Betson (2006) Sample	$44,728	$46,140	$49,834	$48,341
Alcohol and Tobacco				
Complete Sample	2.22%	1.57%	1.48%	1.37%
Partial Sample	2.26%	1.57%	1.50%	1.42%
Betson (2006) Sample	2.20%	1.80%	1.60%	1.40%
Adult Clothing				
Complete Sample	2.75%	2.19%	1.79%	1.60%
Partial Sample	2.49%	2.08%	1.63%	1.52%
Betson (2006) Sample (Apparel)	Not Available			
Food				
Complete Sample	15.88%	15.64%	16.15%	17.54%
Partial Sample	16.30%	16.07%	16.27%	18.17%
Betson (2006) Sample	16.00%	16.70%	17.20%	19.50%
Food at Home				
Complete Sample	10.43%	11.38%	11.96%	13.42%
Partial Sample	10.85%	11.73%	11.99%	13.83%
Betson (2006) Sample	Not Available			

complete data set, while food at home contributes a consistently smaller share to total expenditures in the complete data set as compared to the partial data set. This second observation is to be expected, based on the assumptions being made in this research for the Rothbarth and Engel models. Finally, these discrepancies may very well appear as differences in the final calculations numbers, which will be seen in Sections 6 and 7 of this study.

On a final note, the more aggregately defined variable, food, which encompasses food purchased for home use as well as food purchased for not-at-home consumption, has a rather interesting trend: it would appear as though the average share of expenditures devoted to food in general drops (albeit slightly) when moving from a household with zero children to one with one child. The trend seems to "adjust" itself for households with greater than one child and increases as the number of children increases. This trend does not, however, appear in the Betson (2006) numbers. Additionally, as previously noted, differences in the actual sampling years from the CEX in this study versus Betson (2006) could indeed be the primary reasons for dissimilar trends, and in this case, why the food percentages are smaller in this study's samples as compared to Betson's.

5.4 Concluding Remarks

As noted in the subsections above, we have essentially implemented the same data restrictions as Betson has indicated, with the intention of replicating his results from the 2000 Engel study and the 2006 Rothbarth study, as was stated as being one of the goals of this study. There are some slight deviations from Betson's prescribed list of deletions but these were appropriately justified and moreover, deemed inconsequential to final regression results. A set of statistical descriptions were then presented, in an effort to set the stage for the following section, where we present the results from running the regression equations found in Section 4.

6. REGRESSION RESULTS

The main components for this section include the regression output as well as estimated expenditures on children, based on the regression results. The three models' regression results will be presented first followed by the estimations of expenditures. Discussion of expectations will be included throughout.

6.1 Engel Results

Espenshade (1984) comments on his experience in running several different regression models and proposes a set of guiding criteria to judge the lot; these standards are advantageous as a general benchmark for the results considered here and will be incorporated as well as possible. First of all, Espenshade (1984) echoes the sentiment held by many – that the proxy for household standard of living should most certainly be one that is featured prominently (relative to other potential proxies) in household expenditures, as food and food at home usually are. In either data set, food at home (the dependent variable in our case) ranges from a low 10.4 percent to a high of 13.8 percent (see Table 5.3.2). The reality of this presumption is a reduction in random sampling error. Secondly, Espenshade (1984) distinguishes a so-called "good" regression as one that produces the following results: (i) the estimated coefficients of all the age-sex variables should indeed have the same sign, (ii) the coefficient on total expenditures should have the opposite sign as those found in (i), and (iii) all of the age-sex variables as well as the total expenditures variable should be statistically significant. While Espenshade (1984) created this list specifically for his use in judging the Engel Model, we feel that the criteria presented can be generalized and are applicable to both models. The remainder of this discussion, then, will relate back to these third-party recommendations when judging the results from any of the regressions, Rothbarth included.

Recall that a primary assumption in the Engel model is that as total expenditures rise, the percent of the budget devoted to food should fall (see Section 3.1). For this to be confirmed, we would expect to see a negative coefficient on all total expenditure variables, and in either case, this is true, as can be seen in Table 6.1. Additionally, then, we would expect the opposite sign (positive) on the children variables, which is also the case. The numerical impact on the dependent variable when moving from the youngest age group to the oldest is quite consistent, despite the change in data set: as children get older, the share of expenditures devoted to food appears to steadily increase. The only exception to this can be found in the results from the partial data set (see Table 6.1). The ck35 variable produces a digression from the established trend within the regression results, as it drops in value after first considering the ck02 age group. As the results in Table 6.1 indicate, then, all three of the Espenshade's recommendations introduced at the beginning of this section are met, regardless of data set, with the exception of ck35 variable in the partial data set results, which is not statistically significant.

As an empirical cross-reference we can also review the results from an earlier Betson study (which is, in reality, his most recent study on the Engel model) produced in 2000, wherein Engel estimates are presented on data from the 1996-98 CEX. Indeed, results found in the Betson (2000) study differ substantially with those found in this thesis work, specifically in regards to the age-sex variables. Absolutely none of the age-sex variables come back as statistically significant and furthermore, one of the values is negative (k02 variable). By Espenshade's criteria, Betson's (2000) regression results might indeed qualify as sub-standard, and would be subject to an improved analytical inspection.

Lastly, there is quite a strengthening in the explanatory power attributed to this model vis à vis the Rothbarth model, which furthers the confidence behind the robustness of the Engel approach (see Table 6.3 for results pertaining to the Rothbarth Model). The lower bound of the two results is a rather high 55 percent while the upper bound obtained from the partial data set is an even higher 60 percent. In general, then, we might conclude at this point that this set of regressions appears more reliable and (mostly) indifferent to data set assumptions, although the calculation of child cost shares most certainly could confirm or deny this claim, as we will see in Section 7.

Table 6.1 Estimate of Equation (5): Complete and Partial Data Sets, Engel Model

Dependent Variable: log of total expenditures on food at home shares

Variable	*Coefficient (Complete)*	*Coefficient (Partial)*
LNFSIZE	-0.96141** (0.09022)	-0.96694** (0.11629)
LNPCTEXP	-1.07623** (0.0164)	-1.15038** (0.02136)
LNPCTEXP2	-0.17807** (0.00199)	-0.18505** (0.00253)
CK02	0.34112* (0.15567)	0.44191* (0.1991)
CK35	0.45015** (0.15341)	0.28161 (0.19594)
CK612	0.55482** (0.14068)	0.58785** (0.18167)
CK1315	0.64823** (0.17095)	0.63584** (0.2159)
CK1617	0.66411** (0.15617)	0.67224** (0.19803)
CA1825	-0.44793** (0.05783)	-0.41097** (0.08539)
CA2635	-0.14911** (0.04069)	-0.16014** (0.05142)
CA4655	0.11858** (0.0402)	0.18542** (0.04892)
CA5660	0.0374 (0.05045)	0.08526 (0.06118)
BLACK	-0.15508** (0.03685)	-0.17079** (0.04548)
HD_NO_HS	-0.0119 (0.03669)	-0.00233 (0.04504)
HD_COLL	0.06345** (0.2298)	0.07677** (0.02802)
SP_NO_HS	0.00487 (0.0405)	-0.05929 (0.05149)
SP_COLL	0.05767* (0.02303)	0.04881 (0.02799)
W_WORK	-0.00896 (0.03882)	0.01766 (0.04864)
FTIME	0.00431 (0.02522)	0.03277 (0.03112)
TWOERN	-0.04441 (0.03325)	-0.02863 (0.04134)
Constant	-0.73076** (0.07998)	-0.77744** (0.10021)
N	7,259	4,181
$\overline{R}^2$	0.553	0.5966

Standard errors reported in parentheses. ** indicates significance at the 1% level, and * indicates significance at the 5% level in a two-tailed test.

6.2 Age-Invariant Results

The estimation of the Age-Invariant Model is the result of the regression equation presented in Section 4.2 of this study. Recall that the major difference between this version of the Engel Model and the full Engel Model presented in Section 6.1 is the absence of the child-age proportion variables. In their place we find five kid variables that are included to account for the presence of a child within the household, regardless of age. It directly follows, then, that the coefficients on these variables, as well as those found on the remaining independent variables included in the regression, *cannot* be compared to the full version of the Engel Model. Moreover, because this is essentially the debut of this model in the academic literature[13], it is not possible to compare it with previous work. However, the Espenshade set of criteria (see Section 6.1) are, in fact, relevant here, and the a priori assumptions (see Section 3.1) behind the full Engel model are still in effect here. Thus, a reasonable preliminary assessment is perhaps possible.

Nearly all of the variables appearing in this regression are statistically significant, and more importantly, the child variables and the total expenditures variables maintain the relationship that Espenshade (1984) suggested they have (see Section 6.1), indicating that Engel's assumptions are indeed upheld. Moreover, the numerical impact on the dependent variable (the share of expenditures on food at home) from adding an additional child to the household can be seen very clearly with either data set result: an increase in the number of children is correlated with an increase in the share of the budget devoted to food (see Table 6.2).

The explanatory power from either result is quite comparable to that found using the more detailed Engel model, as shown in Table 6.2. The complete data set results produce an R-square of .54 while the partial data set results produce an R square of .58. The quantitative impact on the estimated child shares of not accounting for children's ages in the regression remains to be seen; it is hypothesized that results will be skewed

[13] While the initial introduction of this model came via a report to the Florida Legislature as noted in Section 1, the regression equation in that report is not exactly the same as this one: there are some additional demographic variables included, such as region of the country, that were omitted here due to our goal of replicating Betson's (2006) study.

upward due to the averaging across ages as well as due to the omission of accounting for differences in age-related food consumption patterns.

Table 6.2 Estimate of Equation (5) : Complete and Partial Data Sets, Age-Invariant Model

Dependent Variable: log of total expenditures on food at home shares

Variable	*Coefficient (Complete)*	*Coefficient (Partial)*
LNPCTEXP	-0.69761**	-0.7554**
	(0.01325)	(0.01731)
LNPCTEXP2	-0.16884**	-0.17571**
	(0.00194)	(0.00249)
KID1	0.15497**	0.17682**
	(0.02566)	(0.03229)
KID2	0.34089**	0.31863**
	(0.02406)	(0.02985)
KID3	0.46915**	0.46272**
	(0.03341)	(0.04067)
KID4	0.51717**	0.56467**
	(0.0578)	(0.07247)
KID5	0.72396**	0.76362**
	(0.12148)	(0.15291)
BLACK	-0.14876**	-0.16166**
	(0.03733)	(0.04639)
HD_NO_HS	-0.02155	-0.02159
	(0.03716)	(0.04594)
HD_COLL	0.06415**	0.07932**
	(0.02324)	(0.02854)
SP_NO_HS	-0.0205	-0.08107
	(0.04102)	(0.05252)
SP_COLL	0.04256	0.03072
	(0.02321)	(0.02846)
W_WORK	0.05335	0.08255
	(0.03899)	(0.04924)
FTIME	-0.000277	0.02471
	(0.02549)	(0.03167)
TWOERN	-0.08416*	-0.07431
	(0.03345)	(0.0419)
Constant	-0.91211**	-0.86368**
	(0.03277)	(0.04009)
N	7,259	4,181
$\bar{R}^2$	0.5411	0.5799

Standard errors reported in parentheses. ** indicates significance at the 1% level, and * indicates significance at the 5% level in a two-tailed test.

6.3 Rothbarth Results

According to the Rothbarth model, if using adult clothing as a proxy for household standard of living, we would expect to see purchases of adult clothing fall as the number of children present in the household rises (see Section 3.2). This is indeed the case when considering the results from using either the complete or the partial data set (see Table 6.3). That is, the coefficients on all of the child variables are negative, implying that, all else constant, an increase in the proportion of children of any age to the household will result in a net decrease in expenditures on adult clothing, and further, a reduction in the standard of living for that particular household.

The numerical impact on the dependent variable when moving from the youngest age group (ck02) to the oldest age group (ck1315), however, is not the same in both models. First of all, one should note that *none* of the children-age variables are statistically significant when using the complete data set, while *all* of the children-age variables are statistically significant when using the partial data set. Furthermore, the benchmark results for this model, the Betson (2006) study of Rothbarth, tend to agree slightly more on this particular account with the results obtained from the complete data set: only the age group of children between 3 to 5 years old is statistically significant in the Betson (2006) study. Second of all, the movement from one age group to the next is not consistent with other findings under the complete data set, while the accepted trends actually can be found under the use of the partial data set. Meaning, when using the complete data set, the youngest age group starts on a negative note, but as one moves to the older age groups, the numerical impact on the dependent variable becomes more positive, until the 13 to 15 year old group, when it becomes more negative again. The significance of this note is that this would suggest that older children require fewer resources than younger children, which is not what the literature suggests. The Betson (2006) study produces a similar movement to that found when using the complete data set. The shape in the child-age proportion results obtained from using the partial data set, however, is more consistent, not just with what we would expect, but also from what has been noted in the literature. Starting on a negative note with the youngest age group, as the age groups increase in age, the net impact on the dependent variable becomes increasingly negative, suggesting that older children tend to require more resources. Because the only regression that produces statistically significant values for the child-age variables also coincides with the expected movement of values, perhaps we should dismiss the other regressions, Betson's (2006) included, on this point, since the child-age values they produce are essentially assumed to be zero. The Espenshade (1984) criteria introduced in Section 6.1 would tend to agree with this conclusion.

A similar variable to the collection of the child-age proportion variables measured in the regression that contains a rather different assumption relative to the child-age variables is the inclusion of the log of family size. Holding *per capita* total expenditures constant, when increasing family size, we would expect to see an increase in total spending on adult clothing; affirming this assumption, in every regression, the coefficient on family size is indeed positive (see Table 6.3). However, only the Betson (2006) study and the regression using the partial data set contain values that are statistically significant. In general, this variable is included to measure how many people reside in a household, ignoring whether or not the additional members to a one-adult minimum household are adults or children. In this study, however, recall that the initial sample pool from which the partial and complete samples were drawn was restricted to only include husband and wife (two adult) households (see Section 5.2). Thus, an increase in family size corresponding to this study's sample is directly linked with the introduction of children to the household. One would expect the coefficient on this variable to be *negative*, since Rothbarth's second assumption requires an inverse relationship between spending on adult clothing and number of kids in a household. In fact, as shown in Table 6.3, in either data set, partial or complete, this parameter estimate is *positive*, which is also the case with Betson's 2006 study. Consequently, this would violate Rothbarth's second assumption (see Section 3.2).

Another assumption central to the Rothbarth model is that adult clothing is not considered a necessity, rather, a luxury good, and thus, would not behave as food would behave when increasing total expenditures. Therefore, when increasing total expenditures, we would expect to see an overall increase in expenditures on adult clothing as well; both the complete and partial regression results support this assumption (see Table 6.3). While in either of these results the log of per capita total expenditures

Table 6.3 Estimate of Equation (6): Complete and Partial Data Sets, Rothbarth Model

Dependent Variable: log of total expenditures on adult clothing		
Variable	*Coefficient (Complete)*	*Coefficient (Partial)*
LNFSIZE	0.18199 (0.12094)	0.56119** (0.16021)
LNPCTEXP	0.66092** (0.02199)	0.72883** (0.02943)
LNPCTEXP2	0.07016** (0.00267)	0.07738** (0.00349)
CK02	-0.2536 (0.20868)	-0.54009* (0.27429)
CK35	-0.17141 (0.20565)	-0.57664* (0.26993)
CK612	-0.16322 (0.18858)	-0.57685* (0.25027)
CK1315	-0.17709 (0.22916)	-0.88518** (0.29743)
CK1617	0.98487** (0.20934)	0.62467* (0.27282)
CA1825	-0.10389 (0.07752)	-0.25832* (0.11764)
CA2635	0.0408 (0.05455)	0.07816 (0.07085)
CA4655	0.00793 (0.05389)	0.06725 (0.0674)
CA5660	-0.0236 (0.06763)	0.10519 (0.08428)
BLACK	0.06939 (0.04939)	0.01497 (0.06265)
HD_NO_HS	-0.08098 (0.04918)	-0.09535 (0.06206)
HD_COLL	0.19259 (0.0308)	0.24346** (0.0386)
SP_NO_HS	-0.05037 (0.0543)	-0.02196 (0.07093)
SP_COLL	0.17763 (0.03087)	0.18083** (0.03857)
W_WORK	0.1196 (0.05203)	0.13453* (0.06701)
FTIME	0.01936 (0.03381)	0.04798 (0.04287)
TWOERN	0.01778 (0.04457)	0.02914 (0.05695)
Constant	5.54752** (0.10721)	5.07618 (0.13806)
N	7,259	4,181
$\overline{R}^2$	0.2023	0.2402

Standard errors reported in parentheses. ** indicates significance at the 1% level, and * indicates significance at the 5% level in a two-tailed test.

and the square of this term are both positive and statistically significant, the Betson (2006) study does not agree with the sign on the squared term. Those results suggest that increases in spending on adult clothing do occur, but at a decreasing rate. Additionally, the value of the squared term is not statistically significant in the Betson (2006) study.

Finally, there are some minor disagreements with all three regression results (the two in this study, and Betson's (2006) study) in terms of the values and statistical significance statuses of the collection of demographic variables, which is to be expected due to inherent sample differences. However, based upon the aforementioned presentation of results, it would appear that the results obtained from the use of the partial data set (see Table 6.3) are more similar to what Betson (2006) finds. Again, this is not surprising, since the exclusion of incomplete responders was a key point in Betson's list of appropriate deletions (see Section 5.2). The R-square affirms this conclusion, as we see an increase from a low 20 percent explanatory power when using the complete data set to a higher 24 percent explanatory power when using the partial data set (see Table 6.3). Betson (2006) reports an R-square of .3258.

6.5 Concluding Remarks

This section has presented the regression results obtained from running the two different data sets that were developed in Section 5 of this study on all three of the models introduced in Section 4 of this study. Some important trends have surfaced and these will surely foreshadow the findings of Section 7, when the regression results are translated into actual child cost share estimates. Some of the trends include the following:

- Both versions of the Engel Model have produced coefficient estimates that are all supported by the a priori assumptions discussed in Section 3.1.
- In addition, the Engel Models are quite consistent with the criteria established by Espenshade (1984) for the assessment of regression models (see Section 6.1).
- The R-square on both the Engel Models is also quite high, regardless of data set used.
- There appears to be a strong trend of consistency across the Engel Models, which will have to be confirmed or denied in Section 7.

- The Rothbarth Model has produced coefficient estimates that are *not* all supported by the a priori assumptions discussed in Section 3.2; indeed, it would appear that the sign on the log of family size term is negative, when it was expected to be positive.
- In addition, the Rothbarth Model varies substantially in terms of either meeting or disagreeing with the Espenshade (1984) criteria presented in Section 6.1.
- The R-square on either Rothbarth regression is quite low, regardless of data set used.
- There appears to be a strong trend of *in*consistency across the Rothbarth Model, which, again, will have to be confirmed or denied in Section 7.
- However, we have come rather close to replicating the Rothbarth results presented in Betson (2006); minor differences may be due to the difference in years sampled or the failure of Betson to disclose all sample restrictions in his study.

We will now move on to Section 7, where the true value of these regression results will be magnified through the translation of the numbers into estimates of actual child share costs

7. CHILD SHARE COST RESULTS

The subsequent step following regression results is transforming these estimations into numerical products that illustrate the resources a child might require under the different models. The average value of all variables was used in the calculations of these numbers. The purpose behind this step is not only to extract a tangible product but also to test, in a sense, each model's sensitivity to the data set being considered. Recall that testing model robustness is one of the primary goals of this thesis (see Section 1). Final numbers as presented in this section indicate that the Rothbarth model is the most sensitive to the underlying data set, as speculated in the previous section.

7. 1 Engel Robustness Results

Table 7.1 contains a summary of the projected share of total expenditures that would be devoted to a child according to age under the results generated by the Engel regression presented in Section 6.1. As clearly indicated by the table, the average

Table 7.1 Test for Robustness: Engel Model			
	Number of Children		
Data Set and Age Group	**1**	**2**	**3**
Age: 0 - 2			
Complete	17.57%	26.06%	32.41%
Partial	20.40%	30.29%	37.63%
Difference	-2.83	-4.23	-5.22
Age: 3 - 5			
Complete	19.90%	29.37%	36.50%
Partial	17.16%	25.78%	31.98%
Difference	2.74	3.59	4.52
Age: 6 -12			
Complete	22.06%	32.33%	40.13%
Partial	23.21%	34.10%	42.25%
Difference	-1.15	-1.77	-2.12
Age: 13 - 15			
Complete	23.92%	34.83%	43.15%
Partial	24.11%	35.29%	43.68%
Difference	-0.19	-0.46	-0.53
Average Difference Between Partial and Complete: 2.45			

difference between calculated shares, accounting for the two different underlying data sets, is a mere 2.45 percentage points. While we have not yet noted the average differences found using the Rothbarth Model, and thus, we can not yet conclude anything about relative differences, in absolute terms, an average difference of 2.45 percentage points is perhaps what one might consider a good indication of consistency. Moreover, when one looks at the individual age groups rather than the aggregate, even smaller differences abound: for example, the oldest age group, as illustrated in Table 7.1, shows near replication. These results show promise that perhaps the Engel Model may be the more robust of the two. That is, it makes little difference if one decides to eliminate observations based on completed interview responses or not, the Engel model will still provide similar estimates.

The largest differences on the other hand, are found in the youngest age groups, where replications of cost share estimations are not nearly as exact. Furthermore, these calculations are also subject to a deviation from trends in regards to which data set tends to quote values on the higher side. Generally speaking, the partial data set is found to generate larger estimates, with the exception coming from the ck35 age group, where the complete data set exhibits larger values. This may be due to the irregularity in the trend found during the regression results (see Section 6.1) under the partial data set output. Recall that it was this age group that fell in value (using the ck02 age group as a staring point), which is not what we would have expected, occurring only in the partial data set results, and producing the only value that was not statistically significant from the child-age collection of variables (see Table 6.1). There may, as a result, exist a small inconsistency with the reported data that is somehow correlated with whether or not the household was a complete responder or not. Table 7.1 summarizes these results.

7.2 Age-Invariant Robustness Results

The Age-Invariant Model is the owner of the smallest differences in estimates due to data set, as presented so far, and found in Table 7.2: only .18 percentage points, on average, separate the two regression results. The almost exact replication of results may very well be due to the omission of the child-age proportion variables that differentiate this version of the Engel Model with the more detailed one. By not accounting for differences in age-related consumption patterns, and instead, averaging the values used in the regression, this act has essentially eliminated most of the differences appearing in the fuller version of this model.

Table 7.2 Test for Robustness: Age-Invariant Model			
	Number of Children		
Data Set	**1**	**2**	**3**
Complete	35.61%	54.52%	62.01%
Partial	35.76%	54.16%	61.99%
Difference	-0.15	0.37	0.02
Average Difference Between Partial and Complete: 0.18			

7.3 Rothbarth Robustness Results

Table 7.3 contains a summary of the projected share of total expenditures that would be devoted to a child according to age under the Rothbarth model. The general trend, regardless of data set, suggests that younger children require the fewest resources, while the older age groups seem to require substantially more. The largest differences in values are found in estimating shares for the youngest age group, ck02, while the closest values are found in the ck612 age range.

On average, the difference in values produced by either data set is 7.49 percentage points. This is by far the largest difference of the three models being tested. Not only is this result relatively large, but in absolute terms, an inconsistency this wide can have severe impacts on recommendations for child support guidelines. This would indicate, then, that the Rothbarth Model indeed is the less robust of the two approaches, as defined by the requirements specified in Section 1 of this study. Furthermore, the complete data set seems to consistently produce larger share values over the partial, with the exception of the oldest age group, ck1315, where the partial data set values are in fact greater. Comparing these final estimates to those found in Betson (2006) confirm the notion that the Rothbarth model is quite sensitive to the underlying data set. For example, Betson (2006) produces results that range from a low 25 percent of expenditures for one child to a high 44 percent of expenditures for three children, while the estimates shown in Table 7.3 range from a low 29 percent to a high 63 percent, which is veritably different.

To locate the source of the driving forces behind the differences within these two results one has to venture back to the regression results. Comparisons indicate that the major differences are found in (i) the value of the constant (our constant is more largely positive relative to the one found in Betson's 2006 study), (ii) the value of the log of per capital total expenditures squared (ours is positive, Betson's (2006) is negative), and (iii) some of the values of the demographic variables (a few vary in sign as compared with Betson's (2006) findings). One would then be inclined to conclude that either Betson failed to make note of all of his sample restrictions, or, in fact, the excluded three years' of data from the CEX make a rather significant difference, enough to produce this wide range of suggested shares.

Table 7.3 Test for Robustness: Rothbarth

	Number of Children		
Data Set and Age Group	**1**	**2**	**3**
Age: 0 - 2			
Complete	34.26%	50.04%	60.09%
Partial	29.42%	42.27%	51.75%
Difference	4.85	7.77	8.34
Age: 3 - 5			
Complete	31.86%	47.27%	57.13%
Partial	30.45%	43.55%	53.18%
Difference	1.41	3.72	3.95
Age: 6 -12			
Complete	31.62%	46.98%	56.83%
Partial	30.46%	43.56%	53.21%
Difference	1.16	3.42	3.62
Age: 13 - 15			
Complete	32.03%	47.47%	57.35%
Partial	38.45%	53.08%	63.43%
Difference	-6.42	-5.61	-6.08
Average Difference Between Partial and Complete: 7.49			

Yet another strong force that may be dictating how these variables act under different data sets is the fact that the Rothbarth Model is based on a share of total expenditures that is quite small, which may introduce measurement errors. For example, it is not quite difficult to imagine the tendency for people to forget the irregular purchases of a non-every day item such as clothing when being interviewed for the CEX, versus an everyday, stable purchase, such as food. Errors in reporting these expenditures will indeed be amplified when further demarcating the data set into child-age related groups, which is what may be happening in Table 7.3.

7.4 Child Share Cost Estimations: All Models

When averaging the estimates across age groups and data sets in the full version of the Engel Model we find that one child would require 20.64 percent of expenditures, two children would require 30.54 percent, and three children would require 37.83 percent of expenditures. Despite an almost exact replication of estimates, and a strong contender for most robust model, the calculated share values under the Age-Invariant Model are, in fact, the highest of the three models considered: one child would require 35.69 percent, two children would require 54.34 percent, and three children would require 62 percent of a family's total expenditures. Finally, averaging the range of estimates under each age group generated by the Rothbarth Model indicates that, under this model, one child would require 32.32 percent of total expenditures, two children would require 46.78 percent, and three children would require 56.62 percent. Table 7.4 summarizes these results.

Table 7.4 Average Share Estimations by Model

	Number of Children		
Model and Data Set	**1**	**2**	**3**
Rothbarth: Complete	32.44%	47.94%	57.85%
Rothbarth: Partial	32.19%	45.61%	55.39%
Average Rothbarth	**32.32%**	**46.78%**	**56.62%**
Engel: Complete	20.86%	30.65%	38.05%
Engel: Partial	21.22%	31.36%	38.89%
Average Engel	**21.04%**	**31.01%**	**38.47%**
Age-Invariant: Complete	35.61%	54.52%	62.01%
Age-Invariant: Partial	35.76%	54.16%	61.99%
Average Age-Invariant	**35.69%**	**54.34%**	**62.00%**

These findings, while characterized by a rather believable notion of what a child might actually require, are not consistent with the literature, with respect to which model is considered an upper bound and which model is considered a lower bound. It is a commonly held impression that the Rothbarth estimates are to be the *lower* bound, while the Engel estimates are to be the *upper* bound; here it is simply not the case when comparing the full version of the Engel Model and the Rothbarth Model. Moreover, the Age-Invariant Model and the full version of the Engel model differ quite markedly with respect to what each suggests the share of expenditures might be for a child, with the Age-Invariant Model generating results that are much larger than those proposed by the full Engel Model. This would appear to suggest that the Engel Model, while mostly indifferent to the underlying data set, is perhaps quite sensitive to model specification. Furthermore, it may very well be, then, that the manner in which these models are constructed will produce biased estimates, and as proposed by some of the papers found in the literature review (Section 2) of this study, the direction of the bias is not necessarily known. If that is indeed the reality, then perhaps the accepted notion of which model "should" be the upper or lower bound may be in need of a reassessment.

In addition, it is important to re-introduce the argument against the Rothbarth case with respect to the discussion on robustness. As mentioned previously, when looking at the results generated by the Rothbarth Model under the isolation of child-age related variables, we see quite a bit of variance in the estimates. However, when aggregated across data set and age groups, the Rothbarth Model seems to perform quite consistently. As shown in Table 7.4, the average share estimation for either the complete or partial data set is fairly similar. This would suggest, then, that on average terms, the Rothbarth Model may not be as sensitive to the data set as noted earlier.

A further deconstruction of these estimates may shed some more light on the matter. By isolating the marginal impact of each child, we can gain yet another angle in how these models may be different or similar. For example, under the full Engel Model, the marginal impact of the second child using the averaged estimates suggests that adding a second child to the household increases expenditures by 47 percent, while the marginal impact of a third child is roughly around 24 percent (see Table 7.4.1). While these values are quite a bit higher than the marginal impact results attributed to the Rothbarth Model, and in this case, could be considered for the title of upper bound, the Engel Model marginal impact results are in line with what is produced under the per capita model. The marginal impact of the third child is a bit high, regardless of the benchmark chosen; recalling the underlying calculations, one can see that this is where the largest differences due to data set are produced, and this may be a potential cause.

In contrast, the Age-Invariant Model suggests that the marginal impact of the second child is right at 50 percent, while the marginal impact of the third child is around

Table 7.4.1 Marginal Impact by Model

	Child	
Model	**2^{nd}**	**3^{rd}**
Rothbarth	45%	21%
Engel	47%	24%
Age-Invariant	52%	14%

15 percent. While this second set of estimates appears high in comparison with those found under the previously considered model, they are rather similar to those found using the per capita method. This is most likely due to the non-separation of children by age group; indeed, under the Age-Invariant Model the age of the children is assumed irrelevant to changes in expenditures, and thus, expenditures would be essentially averaged across ages for each additional child. While this greatly simplifies the model, and appears to make it more indifferent to the underlying data set, it may be a gross oversimplification that results in inflated estimates. For example, the cost share estimate ranges for one child generated by the more detailed Engel model starts with a low 17.16 (ck02) and reaches a high 24.11 (ck1315) (see Table 7.1). For two children this range expands to an almost 10 percentage point difference. This suggests that not taking into account the children's ages within the regression and instead, using their average ages, results in more elevated estimates. This tradeoff may be acceptable, if one is most concerned with model robustness.

Finally, the marginal impact of a second child in the Rothbarth Model based on the averages found in Table 7.3 suggests that adding a second child to the family would result in an increase of around 45 percent, while the marginal impact of adding a third child is around 21 percent. These numbers are actually rather consistent with the findings in Betson (2006).

7.5 Concluding Remarks

We have discovered in this section that the based on the results for testing for robustness introduced in Section 7.1, the Rothbarth Model appears to be the least robust of the three models presented, before averaging across data set and age group. Furthermore, the Age-Invariant Model appears to be the most robust of the three models. Cost share estimates produced by the three models suggest that the full version of the Engel Model is in fact a lower bound, the Rothbarth Model estimates fall somewhere in the middle, while the Age-Invariant Model produces the largest cost share estimates. This large difference in estimates produced by the Engel approaches is assumed to be due to sensitivity to model specification on the part of the Engel approach. We will now move on to the final section of this study, where we summarize all findings and present future opportunities for research.

8. CONCLUSIONS AND FUTURE WORK

The purpose of this study was to further investigate three approaches to the empirical estimation of child expenditures: the Rothbarth Model, the Engel Model, and the Age-Invariant Model, and secondly, to attempt a discovery of any sensitivities that might render a model weaker than its counterparts. More specifically, the intention was first to replicate the results found in the Rothbarth estimations produced by Betson (2006), secondly, to test for model robustness, and thirdly, to present the respective child cost share estimations produced from each model's regression results.

The regression results presented in Section 6 indicate that the regression that utilized the partial data set came rather close to replicating the results found in Betson's (2006) Rothbarth study. Differences appeared to be mostly isolated to the constant term, the change in sign on the squared term for total expenditures, as well as a few of the demographic variables. This suggests that the three years of excluded data from the CEX in this study may have had an impact; another possibility may be the failure of Betson to disclose all data restrictions imposed in his 2006 study.

Based on the results presented in Section 7, the Rothbarth Model had the highest differences among estimates when considering the impact of the children based on ages, suggesting that it is the model that is most sensitive to the underlying data set in this case. This would most certainly lend credence to those Engel supporters that believe basing child support payments on a model whose foundation is a proxy for wellbeing that just barely accounts for 4 percent of total expenditures is, in fact, inappropriate, and perhaps even justify their almost 13 percent of total expenditures proxy as a much better candidate. Results from the Engel Model show a reliability lacking in the Rothbarth; moreover, despite the Age-Invariant Model producing estimates that are larger than its more detailed associate, it revealed an even stronger indifference to a particular data set. This large difference, however, may suggest that the Engel approach is quite sensitive to model specification, despite an apparent indifference to the data set being used.

Trends of the behavior of all three models alluded to increasing costs as children age, in general, as well as a smaller average cost per child, as the number of children in a household increase. However, the cost share estimations produced within each model were certainly not all uniform; there was a wide range of values presented that only serve to further the argument that the answering of the question, 'What is the cost of a child?' is not nearly straightforward in any respect. Furthermore, there are some substantial disagreements regarding the marginal impact of the second and third child. For example, the Age-Invariant Model generated a value of 52 percent for the impact of a second child, while the Rothbarth purports this value to be closer to 45 percent (see Table 7.4.1). Likewise, there are also large differences between the two versions of the Engel Model, with regard to the marginal impact of the third child. These disparate estimations trace their incongruence back to the original regression results and the model specification, and thus, only serve to further highlight the great difficulty it is to settle upon agreeable estimates.

The observations above lead our discussion to the actual average estimations that were generated from the regression results. Based on Betson's collection of studies (2000, 2006), one might be persuaded to conclude that output from the Engel Model should not only be considered an upper bound, but should, furthermore, be disregarded, while output from the Rothbarth Model should be considered a lower bound and final numbers upheld. However, the results of this study suggest otherwise. Indeed, there is evidence to believe that the Rothbarth Model is perhaps the weaker of the two, since results indicated a higher sensitivity to the underlying data set. Because this study is essentially an outgrowth from Betson's 2006 study, in an effort to satisfy some lingering questions future research should include running the same regressions on the same years of CEX data utilized in Betson's (2006) study. While replication of his Rothbarth study was close (solely in terms of regression coefficients), the numbers differ enough to affect child expenditure estimations immensely. As Table 8.1 shows, the differences in estimated child expenditure shares range from a low of 7 percentage points to a high of 13 percentage points; it is inferred that the missing data is what is driving these differences.

The inclusion of the Age-Invariant Model served mostly to magnify the hypothesized belief that the Engel Model is the more robust of the models considered; however, future work should also consider an "Age-Invariant" Rothbarth Model that contains the same modifications implemented in the simpler Engel version. This will help develop a more level platform with which to compare these models, and perhaps allow us to gain further insight into the impact of not accounting for a child's age within a regression equation.

Table 8.1: Estimated Shares by Model			
	Number of Children		
Model	1	2	3
Per Capita	33%	50%	60%
Betson (2006): Rothbarth	25%	37%	44%
Arce-Trigatti: Rothbarth	32%	47%	57%
Betson (2000): Engel	30%	44%	52%
Arce-Trigatti: Engel	21%	31%	39%
Arce-Trigatti: Age-Invariant	36%	54%	62%

Finally, the sum of all these observations should certainly lead one to ponder: which results should be deemed "correct"? While no other study in the literature lays claim to having uncovered the 'true' cost of children, this study, unfortunately, is no different. It is quite difficult (in fact, currently impossible) to accurately measure just how far off the mark this, or any of these studies, are from the 'true' cost, due to the lack of appropriate individual-cost measuring tools applicable to a household. Thus, we can only propose, recommend, and theorize just what these numbers might actually look like. In terms of which is most correct, based on the discussions above, this particular study would stand behind the Engel Model as a more truthful and reliable child cost indicator.

9. BIBLIOGRAPHY

Barten, A.P. (1964), "Family Composition, Prices and Expenditure Patterns." P. Hart, G. Mills, and J.K. Whittaker (eds.), *Econometric Analysis for National Economic Planning*, 16th Symposium of the Calston Society, London: Butterworths.

Bassi, Laurie and Burt Barnow (1993), "Expenditures on Children and Child Support Guidelines." *Journal of Policy Analysis and Management*, Vol. 12, No. 3, 478-497.

Betson, David (1990), "Alternative Estimates of the Cost of Children from the 1980-86 Consumer Expenditure Survey." (Report to U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation). Institute for Research on Poverty, University of Wisconsin-Madison, Special Report Series No. 51.

Betson, David (2000), "Parental Expenditures on Children: A Preliminary Report." Unpublished manuscript, University of Notre Dame, Department of Economics. Retrieved July 27, 2008, from http://www.nd.edu/~dbetson/research/documents/CostofChildren2000.pdf

Betson, David (2006), "Parental Expenditures on Children: Rothbarth Estimates." A report prepared for Policy Studies, Inc., for the State of Oregon. Retrieved July 27, 2008, from http://www.nd.edu/~dbetson/research/documents/CostofChildren2006.pdf

Betson, David, Eirik Evenhouse, Siobhan Reilly, Eugene Smolensky (1992), "Trade-Offs Implicit in Child-Support Guidelines." *Journal of Policy Analysis and Management*, Vol. 11, No. 1, 1-20.

Deaton, Angus and John Muellbauer (1986), "On Measuring Child Costs: With Application to Poor Countries." *The Journal of Political Economy*, Vol. 94, No. 4, 720-744.

Deaton, Angus and Christina Paxson (1998), "Economies of Scale, Household Size, and the Demand for Food." *The Journal of Political Economy*, Vol. 106, No. 5, 897-930.

Engel, Ernst (1857), "Die Productions und Consumtionsverhaltnisse des Konigsreichs Sachsen." *Zeitschrift des Statiscshen Bureaus des Koniglich Sachishen Ministeriums des Innern*.

Espenshade, Thomas J. (1973), *The Cost of Children in Urban United States*, Population Monograph Series, No. 14, University of California, Berkeley: Institute of International Studies.

Espenshade, Thomas J. (1984), *Investing in Children: New Estimates of Parental Expenditures*, Washington, DC: Urban Institute Press.

Garfinkel, Irwin (1992), *Assuring Child Support*, New York: Russell Sage Foundation.

Lazear, Edward P. and Robert T. Michael (1988), Allocation of Income within the Household, Chicago: University of Chicago Press.

Lino, Mark (2001), "USDA's Expenditures on Children by Families Project: Uses and Changes Over Time." *Family Economics and Nutrition Review*, Vol. 13, No. 1, 81-86.

Lino, Mark (2007). "Expenditures on Children by Families: 2007 Annual Report." U.S. Department of Agriculture, Center for Nutrition and Policy Promotion. (Miscellaneous Publication No. 1528-2007). Retrieved July 27, 2008, from http://www.cnpp.usda.gov/Publications/CRC/crc2007.pdf

McCaleb, Thomas, David Macpherson, and Stefan Norrbin (2004), "Review and Update of Florida's Child Support Guidelines." Report to the Florida Legislature, Tallahassee, Florida: Florida State University, Department of Economics.

Morgan, Laura (2004), "The Economics of Child Support Guidelines: A Short Examination of the 'Cost Shares' Model." *American Journal of Family Law*, Vol. 18, 149-154.

Nelson, Julie A. (1986), "The 'Cost of Children' Question: Human, Capital, Fertility or Welfare." Mimeo, University of California, Davis, Department of Economics (1986).

Nelson, Julie A. (1992), "Methods of Estimating Household Equivalence Scales: An Empirical Investigation." *Review of Income and Wealth*, Series 38, No. 3, 295-310.

Nelson, Julie A. (1993), "Household Equivalence Scales: Theory versus Policy?" *Journal of Labor Economics*, Vol. 11, No. 3, 471-493.

Percival, Richard and Ann Harding (2005), "The Estimated Cost of Children in Australian Families in 2005-06." Commissioned research report for the Ministerial Task Force on Child Support, National Centre for Social and Economic Modelling, University of Canberra.

Pirog-Good, Maureen and Patricia Brown (1996). "Accuracy and Ambiguity in the Application of State Child Support Guidelines." *Family Relations*, Vol. 45, No. 1, 3-10.

Rothbarth, Erwin (1943), "Notes on a Method of Determining Equivalent Income for Families of Different Composition." Appendix 4 in Charles Madge (ed.), *War-Time Pattern of Spending and Saving*, National Institute for Economic and Social Research, Cambridge: Cambridge University Press.

Watts, Harold W. (1967), "The Iso-Prop Index: An Approach to the Determination of Differential Poverty Income Thresholds." *Journal of Human Resources*, Vol. 2, No. 1, 3-18.

Williams, Robert G. (1987), "Development of Guidelines for Child Support Orders: Final Report." Washington, DC: U.S. Department of Health and Human Services, Office of Child Support Enforcement.

White, K. R. and R. T. Stone (1976), "A Study of Alimony and Child Support Rulings with Some Recommendations." *Family Law Quarterly*, Vol. 10, 75-91.

Yee, L. (1979), "What Really Happens in Child Support Award Cases: An Empirical Study of Establishment and Enforcement of Child Support Orders in the Denver District Court." *Denver Law Journal,* Vol. 57, 21-68.

BIOGRAPHICAL SKETCH

EDUCATION

Florida State University, Fall 2006 – Summer 2008 (expected)

- **M.S., Applied Economics, 3.86 G.P.A.**

Florida State University, Fall 1999 – Fall 2004

- **B.S., Business: General Management, 3.9 G.P.A.**
- **B.A., Music (Anthropology Minor), 3.9 G.P.A.**

Honors

- Graduated summa cum laude (undergraduate)
- International Study Abroad Program, Valencia, Spain, Summer 2004
- Florida Bright Futures Scholarship, 1999-2003
- Pi Kappa Lambda, Music Honor Society, Inducted 2003
- Florida State University Incentive Scholarship, 1999-2001

Paper

- Arce-Trigatti, M.P. and P. Arce, "Parallel Between Active Learning and Coaching Team Sport Techniques: Analysis and Selected Examples," Annual Conference Proceedings, American Society for Engineering Education, ASEE (2000).[Paper selected for the 2000 NSF-SUCCEED Award for Innovative Teaching, COE, Tallahassee, Florida and profiled in the ASEE PRISM Magazine].

WORK EXPERIENCE

Graduate Research Assistant, Fall 2007-Summer 2008
Florida State University, Tallahassee, Florida

West Coast Production Director, April 2007-June 2007
Bob Knight Photo, Los Angeles, California

Teaching Assistant (online classes), Fall 2006-Spring 2007
Florida State University, Tallahassee, Florida

Production Team Leader, Summer 2001 – December 2005 (Summer/Winter Seasons)
Bob Knight Photo, Tallahassee, Florida

Customer Service Representative, Summer 1999 – December 2005 (Part-Time)
Bob Knight Photo, Tallahassee, Florida

Peer Advisor, Fall 2000 – Spring 2004
Board of Advisors, FSU School of Music, Tallahassee, Florida

COACHING EXPERIENCE

Head Coach, Summer 2007-Spring 2008
U-14 Dragons Girls Travel Team, Top of Florida Soccer Club, Tallahassee, Florida

Co-Head Coach, Summer 2006 – Spring 2007
U-18 Dragons Girls Travel Team, Top of Florida Soccer Club, Tallahassee, Florida

Head Coach, Fall 2006 – January 2007
Girls Junior Varsity, Leon High School, Tallahassee, Florida