The Rothbarth Internal Allocation Model Re-examined:

Semi-Nonparametric and Parametic Tests of Child Gender Discrimination*

Feridoon Koohi-Kamali

Economics Department, City College, CUNY, and SIPA, Columbia University 2019

*****The author wishes to express his gratitude to Hashem Dezhbakhsh, Kevin Foster, David Jacho-Chavez, Esfandiar Maasoumi, and especially to Caterina Ruggeri Laderchi for comments on an earlier draft of this paper; he also thanks Ran Liu and Yibin Liu for their research assistance.

The Rothbarth Internal Allocation Model Re-examined:

Semi-Nonparametric and Parametic Tests of Child Gender Discrimination

Summary

Tests of child gender bias by a model of expenditure on goods solely consumed by adults, such as tobacco, have consistently failed to reveal any evidence of discrimination in countries with severe records of welfare disparity against girls. We argue that son preference and child gender bias is likely to result, in the absence of pre-natal technology, in an increase in the number of children, and propose a gender specification based on the number rather than the age of children in the household as in most previous studies. We also asses the child gender misspecification with another principal source of misspecification, namely total expenditure, by testing parametric specifications of the adult goods Engel curves against semiparametric specifications. We control for endogeneity of both total expenditure and the number of children the estimated budget share Engel curves. Applied to an Ethiopian Expenditure Survey for 2004-2005, we report evidence of discrimination against girls present, in various degrees, in all four pure adult goods available from the survey; although our semiparametric specification tests reveal important functional form misspecifications, the gender effects on consumption remain robust to the functional form specification of the lead expenditure variable; highlighting the pivotal effect of defining gender in relation to the number of children.

Keywords: Intrahousehold Inequality, Adult Goods, Engel Curve, Child Gender, Nonparametric, Semiparametric.

JEL Codes: C01, C14, D12, D13, J16, O55

1-Introduction

Child gender discrimination has long-term consequences for the mortality, health, and education of children. Much evidence is available on welfare disparities by child gender, yet very little research exists on the intrahousehold discrimination mechanism that generates such consequences. This is not for the lack of an effective approach. In particular, Rothbarth (1943) suggested that variations in adults' consumption levels provide a measure of the effect of children on their parents' standard of living, see Deaton and Muellbauer (1986) for its assessment. Deaton (1987, 1989) argued that this approach is promising for analyzing inequality in intrahousehold resource allocation that may be affected by child gender. The approach estimates the differences in the effects of boys and girls on household expenditure for goods not consumed by children, adult goods such as tobacco or adult clothes. The differences in the size and significance of the gender effects provide evidence on the parental preferences for the gender of children. However, despite many applications of the Rothbarth approach, there has been no convincing evidence from consumption patterns that identifies discrimination, leading Deaton (1997) to conclude that "It is a puzzle that expenditure patterns so consistently fail to show strong gender effects even when measures of outcomes show differences between boys and girls."[1] An explanation for this

1 Deaton (1997) contains references on the countries studied to that date, and China is first examined in Burgess and Zhuang (2003); the third attempt on Indian data is Case and Deaton (2002)., while Gibson and Rozelle (2004) apply the model to Papua New Guinea, and Kebede (2008) to Ethiopia. Nearly all the studies employ the *almost ideal* flexible functional form of Deaton and Muellauer (1980) for parametric specification. Of particular interest to this paper are the two semiparametric studies of Bhalotra and Attfield (1998), and Gong et al. (2005), these are the second attempts with, respectively, rural Pakistani and rural Chinese data. This list is unlikely to be

outcome is missing despite the advances that the collective bargaining approach has made to highlight the influence of adult gender on household internal resource allocation. This paper explains this discrepancy by examining parametric and non-semiparametric sources of misspecification in the Engel curves that have been employed with applications of the Rothbarth approach. Moreover, designing effective anti-discrimination public policies requires establishing whether the main cause of gender disparity in welfare outcomes is parental discrimination or budget constrain influences. Such evidence must be obtained from an explicit model of internal inequality, since child gender differences in health or education are also compatible with non-discriminatory internal allocation behavior, see Jensen (2002).[2]

In particular, we show that misspecification of the gender demographic variables is a problem within the existing studies, irrespective of whether they use free or flexible functional forms for expenditure. We examine an alternative to the age-based gender demographic that uses the *number* of children: parents with strong son-preference are likely to achieve the desired gender composition of children through higher fertility when gender selection technology is absent. We also examine a second source of misspecification relates to the flexible almost ideal functional form for the lead expenditure term employed in most

---

exhaustive; the collective bargaining-based extensions of the Rothbarth approach are examined below.

2 Jensen (2002), discussed below, maintains that aggregate male/female sex ratio below its biological value cannot be taken as solely due to gender bias because son preference results in more girls in households with a larger number of children; consequently, division of fixed resources among more children results in lower per capita expenditure on girls, even without unequal internal resource allocation between boys and girls. Another example is the model of Becker and Tomes (1976) and Becker (1981) internal allocation model is based on equal concern and “child-neutral” preferences; parents may invest differently in the human capital of children in accordance with the differences in their genetic endowments.

applications of the model to child gender bias. We examine our Engel curve questions by employing a semiparametric estimator that allows for an unknown functional form of total expenditure in obtaining estimates for the child gender effects. The semiparametric model adopted in this study thus provides an approach that permits combining solutions to parametric misspecification of demographics with nonparametric specification of total expenditure. The usual problem that total expenditure to susceptible measurement error becomes more important in a predominantly rural economy examined in this paper as the share of consumption not going through the market is substantial. Our semiparametric estimator also considers endogeneity for total expenditure and the number of children by the applications of control approach method.

We demonstrate the above approach to the Rothbarth model with an application to data from a 2004-2005 expenditure survey from Ethiopia, a country with severe records of child gender disparities in health and education. Koohi-Kamali (2008) employed the same Ethiopian survey used in the current study, and reports the ratio of primary school enrollment of girls to boys as 0.806, similar to Pakistan (Aslam and Kingdon 2008, table 2); and the population sex ratio of female to male as 0.895, considerably outside the 0.93-0.96 range of naturally occurring ratio (Parazzini et al. 1998). The latter is taken as a broad indicator of female health and mortality, Sen (1990). Yet another indicator is nutritional outcomes for girls. Young Lives (2014, tables 1 and 3) reports food insecurity in Ethiopia in 2009 as 34 percent for boys and 41 percent for girls; while chronic malnutrition based on height-for-age (stunting) in 2013 stood at 27.1 percent for males and 30.9 percent for females. Despite the evidence of gender disparities in welfare outcomes, there is no matching evidence of discrimination in household internal resource allocation that can explain such outcomes. This paper attempts to provide an explanation for that outcome demonstrated with an application of the Rothbarth model to Ethiopia.

The Ethiopian data set employed in this study covers the entire country and provides extensive expenditure items, but not income information. In order to retain consistency with the assumptions of the Rothbarth model, we employ only the data relating to nuclear household observations. Four plausible adult goods are available from this data set: tobacco, coffee, adult clothes, and adult consumption items and services, e.g., for shaving. Results show evidence of statistically significant discrimination against girls in consumption patterns of all four adult goods sampled in the survey after testing for functional form specification of expenditure and the number of children, and after controlling for the endogeneity of total expenditure and the number of children. However, although parametric gender effects remain statistically significant for all four adult goods of this study, for two out of four commodities examined, the functional form specification tests reject the quadratic Engel curves; that said, the gender effect results remain robust to expenditure specification, hence providing more robust evidence of child gender bias. In addition, the econometric controls for potential endogeneity of the logarithm of total expenditure and the number of children appear insignificant in all of the Engel curves examined.

The contributions of this paper consist of providing empirical evidence for the specification of gender child effects on the Rothbarth-based Engel curves; hence demonstrating the pivotal role of the proportion of girls/household size indicator in uncovering evidence of discrimination that goes beyond those documented by Koohi-Kamali (2008) by addressing *a.* the functional form robustness of the lead expenditure variable with respect to child gender effects on consumption, and *b.* control for total expenditure endogeneity that may otherwise result in biased estimates of those effects.

Section 2 discusses the issues related to the Rothbarth model and the requirements for its effective application; section 3 examines the specifications adopted for the semiparametric and parametric models; section 4 deals with data and samples employed; section 5 presents the results inclusive of controls for endogeneity. A final section sums up the findings.

2-Literature on inferring child gender bias from adults' consumption patterns

There are three ways that children affect parental consumption: 1) by changing their tastes, 2) by altering the relative shadow prices of adult vs. non-adult goods, and 3) by lowering resources available for parents' consumption. Rothbarth (1943) suggested a method to isolate the last effect from the first two by employing a pre-selected group of goods that are not consumed by children (e.g., alcohol), so that the addition of a child affects only the parental living standard (i.e., the share of household income going to parents' consumption) but will have little substitution effect. The Rothbarth adult equivalence scale, or cost of a child, is measured by the amount of income required to raise the level of expenditure on adult goods to the level observed prior to a child's arrival, thus providing the basis for a model of intrahousehold resource allocation between adults and children.[3] Crucial for the validity of this method is that the changes in the levels of adult goods consumption must be due solely to changes in income; that is, the demographic change does not additionally increase adult goods consumption through changes in the relative "prices" of non-adult

3 Note the model infers child welfare indirectly by focusing on adults' consumption as an indicator of internal resource allocation.

compared to adult goods. Section 2.1 examines the evidence for the plausibility of this assumption.

2.1 Rothbarth measurement of the cost of children

The models of internal household allocation that presuppose goods that are consumed by a subunit of the household fall into two groups, non-bargaining treating all adults as a subunit, and bargaining-based treating each adult member as a subunit. The former dates back to Rothbarth (1943) who proposed expenditure on goods that reduce the adults' share of household resources as the number of children increase, without having a substitution effects, can identify the effects of children on the adults living standard. Regarding the Rothbarth scales, Gorman (1976) suggested the presence of substitution effects if children lower the relative price of adult goods (e.g., due to an increase in the cost of a child-minder), resulting in higher consumption of such goods. Since the model measures the compensation needed for household income to reach the level attained prior to the addition of a child, its application produces estimated child costs that are too small. Deaton and Muellbauer (1986) examine this issue at some length and conclude that because household consumption in developing countries tends to be closer to subsistence, the Rothbarth underestimation of child costs is unlikely to be significant and more likely to result in estimates that are "appropriate and defensible." Deaton et al. (1989) tests the separability assumption extensively and finds evidence that broadly supports the assumption. Gronau (1988) also argues that because parents may derive utility from their children's consumption, children may reduce the marginal propensity to consumption of adult goods, leading to an additional source of underestimation of child costs. Gronau (1991) tests this notion using US consumption data and reports evidence in support of this stronger version of separability. Barten (1964) proposed a model that accounts for differences in household demographics to affect the relative price of private goods due to differences in the scope of household-level shared public goods. Nelson (1992) has argued that the presence of economies of scale in consumption that are available to larger households (e.g., from housing), may result in both price and income effects in the Rothbarth model. Barten-Gorman scales imply household size-related economies affect consumption of household-level public goods differently from those of household-level private goods (e.g., housing *vs.* food). This issue is examined in an important paper by Nelson (1988) that anticipates the more recent debates (see Deaton and Paxon, 1998); but the evidence obtained so far by non-bargaining models for the Rothbarth scales is based on Engel curves with approximately linear functional forms, e.g., the almost ideal type of Deaton and Muellbauer (1980) employed in Bradbury (1994), for example. As Blackorbry and Donaldson (1993) have argued, the almost ideal model cannot identify equivalence scales because that model's functional form is log-linear in utility (or income); see also Banks et al. (1997), among others, for its non-linear extensions to deal with this problem.

The bargaining-based approach to household internal allocation presupposes existence of goods that are assignable to individual household members, for instance female cloths, Browing et. al. (1994) proposed a Pareto efficient bargaining model of resource allocation between childless men and women that extensively rejects the income pooling hypothesis of internal allocation compared to an alternative based on each adult member's *resource share* of the household income. A number of more recent studies have extended the scope of the resource share model. Browing et. al. (2013) developed a model to demonstrate how each household member's share of total household resources is affected by the benefits of household-level consumption economies. The model is based on gender-specific adult goods, typically adult male and female cloths, and generalizes Barten (1964) and Gorman (1976) by taking into account the impact of household-level public goods on the gender-specific share of household resources incurred by *childless* man and woman.

Another critical advance of the collective model approach has been to consider the impact of children and child welfare on person-specific household resource shares. The initial attempt in Browing and Lechene (2001) analyzed the effect of differences in man and woman resource shares on child welfare by modeling children as a joint consumption, public good by abandoning the efficiency assumption of the collective model and relying on non-cooperative structure that require more complicated theoretical structure and many years of household survey data to estimate a much large number of parameters; some still with imprecise estimation. Bargain and Danni (2011) model children with their own person-specific share of household resources shares and identify resource share differences affected by bargaining factors and consumption economies based on the assumption of unchanged consumption preference of couples with children and childless couples. Dunbar et. al. (2013) has extended that study through two hypotheses, first based on the assumption that household members allocate the consumption budget shar of the same private goods similarly regardless of the number of children; the second that that household members with a given number of children allocate the budget shares of different private goods similarly. The first is intended to identify the person-specific resource share and the test results confirm identification. The test outcomes regarding the assumption public goods inclusive levels of resource shares for each household size, however, produce mixed results that are also considerably less accurate that those based on the first hypothesis. Moreover, the results regarding the impact of the change in the level of child resource share on male and female adult shares are not robust[4],

2.2 Rothbarth measurement of child gender bias

Similar non-bargaining and bargaining-based approaches have also been advocated, though only a few with regard to the gender of children. Testing the Rothbarth model for child gender bias is simpler than the estimation of adult equivalence income, since that task does not focus on estimating the amount of income required to compensate the family for the presence of children; instead it compares, for otherwise similar households, the demographic effects of boys with those of girls in terms of adult consumption expenditure. Unless the presence of possible substitution effects are (child) gender related in important ways, such possible effects would be equally present in adult goods expenditure patterns for both boy- and girl-dominant families; thus, provided a plausible specification is adopted for the gender demographic variables, they cancel out in comparison.[5] Accordingly, the Rothbarth model remains rather an attractive method for identification of child gender effects on consumption patterns, since the focus of analysis is on child welfare or *inequality among children*. Further, the model is well understood, and there is broad consensus on its effectiveness for the estimation of child cost. Gronau (1988), for example, provides a cogent critique of equivalence scales models, and regards the Rothbarth model as "the only" theoretically consistent model among them. Finally, since the model's predictions

---

[4] Regarding reported estimates on the diverted impact of falling share of resource levels for children as child number increases, on man and woman private consumption levels, Dunbar et. al. (2013, 459) state that "whether the [fall in share is] diverted to men or women is specification dependent", that is conditional on the list of included demographic variables.

[5] In the absence of son preference, consumption economies may also affect child gender allocation if more girls live in larger households; the issue is discussed below.

involve no price effects and can hence be tested on a single cross-section budget survey, it has a relatively modest data requirement.

More recent bargaining-based advances in intrahousehold analysis with respect to children have yet to model fully child gender bias effects on adult consumption and provide empirical evidence consistent with the collective approaches. In particular, we note that the more effective one of the Dunbar et. al. (2013) two hypotheses above based on the assumption of unchanged person-specific allocation of a private assignable good is hypothesized with reference to the number of children, hence the test results and estimation remain silent with regard to child gender. Testing for unchanged allocation with reference to child gender is critical for accessing the assumption of constant preference with household demographic change that is built, as in this study, on the basis of invariance to household size, that is, to the number of children. Moreover, Dunbar et. *al.* (2013) test of the second hypothesis for child gender effects across different household sizes for constancy of consumption patterns is equally critical since most types of economies in consumption are related to an increase in household size. Although the study controls for the ratio of girls, the authors cannot infer gender bias from the reported significantly negative child gender effect for Malawi.[6] Similarly, Bargain and Donni (2011) report significantly lower resource shares for girls than boys but maintain that no differences in welfare levels can be inferred from those results.[7] If child gender bias expresses itself in changes in the number of children, then both of the hypotheses proposed in Dunbar et. al.

[6] “We do not claim that these allocations are necessarily unfair or imply inequality in welfare” (p.441); the paper attributes the evidence to the possibility of different metabolic needs of boys and girls.

[7] “The result simply says that what the parents spend on a girl is lower than what they spend on a boy” (p.804).

(2013) are required to extend the tests to unchanged person-specific resource share levels regardless of child number *and* child gender. Perhaps a single hypothesis has potential for identification of both household size driven economies in consumption parameters and for gender driven number of children parameters with son-preference fertility behavior; we are not aware of such an approach. However, it is hard to access the contribution of the bargaining-based approach to modeling child gender effects because such models have so far offered no evidence that son preference affect man and woman living standard in similar but opposite directions, and hence undetectable in the non-bargaining approach; nor has an argument been put forward that child gender evidence against girls represents solely the impact of economies in consumption. In the absence of such evidence, the question posed in Deaton (1997) remains pertinent: why the model has so far failed to uncover evidence of child gender effects in application to the countries with strong records of welfare outcomes favoring boys? The main contribution of this paper is to provide an explanation of the puzzling failure of that model to reveal evidence of child gender in many repeated applications by drawing attention to the importance of specification of child gender bias in relation to the number of children.

We would argue that the misspecification of the child gender variables has been an important factor; almost all previous studies of child gender effects on the Rothbarth model Engel curves employ a functional form for an index of child gender based on proportion of boys and girls in different age groups of children. By assuming gender bias is related to child age, the earlier applications have ignored the systematic demographic evidence that such bias is related to the *number* of children. There is considerable demographic evidence that strong parental child gender preference is highly correlated with the number of children. Parents who desire more male offspring but fail to achieve their target are likely to have more children in order to achieve the target number of boys. Similarly, unwanted girls are more likely to live in larger households and be among the older children (see, for example, Das Gupta, 1987, Muhuri and Preston, 1991, and Clark, 2000). Evidence about the presence of such a mechanism has so far come from demographers, usually in the form of correlations between the proportions of boys and girls. However, pro-boy adult preferences do not always result in additional births because the fertility-stopping rule is unlikely to be homogenous across couples. For some couples, the disadvantage of an additional female birth may outweigh the benefits of an additional male birth. Such couples are likely to stop at a lower parity; even though that decision would be based on son preference.[8] In an interesting demographic study, Jensen (2002) maintains that a biologically unbalanced aggregate sex ratio when son preference is dominant in the population can also be consistent with equal treatment of boys and girls in parental internal resource allocation resource shares as more girls residing in larger households face lower per capita share in the household even if treated as equally as boys. An important implicit assumption in drawing this conclusion is that the fall in per capita income would always lead to a fall in resource share devoted to girls in the absence of unequal child gender allocation. That conclusion would not necessarily follow because an increase in household size also opens up the scope for economies of scale in consumption of public, shared goods, for instance in housing; hence releasing resources for private consumption of each individual household members.

8 Suppose a couple with two girls and no boy most prefers to have two boys and one girl. The couple might regard (an approximate) 50% risk of a third girl as sufficiently high to stop having children.

Jensen does not take fully into account that son preference would also leads to lower birth parity; this is a particularly relevant to the data set employed in this study, explicitly organized to remove any effect of pre-natal technology on the gender composition of children.[9] This study shows the limits of a demographic approach to child gender analysis in terms of the impact of child gender on fertility because, as pointed out by Leung (1991), the correlations evidence can reveal the presence of child gender bias, but *not* its direction; a more plausible approach for obtaining evidence for the latter would be through an examination of household internal allocation *conditional* on gender demographics

Son preference manifests itself through household size, or equivalently the number of children in a nuclear family, and since changes in household size broadly define the scope of economies of scale in consumption, the specification of child gender effect on adult consumption and living standard by this variable requires controlling separately for household size. We therefore examine an alternative to the Deaton gender specification based on the number of children by separating the child gender effects on consumption into a set of child age variables, and proportions of boys or girls in the household; controlling for log of household size separately. A relatively lower expenditure on adult goods for boys would be evidence of bias against girls; that is, the proportion of girls exerts a *positive* effect on adult goods budget shares (see section 3). Koohi-Kamali (2008) applied the Rothbarth model with child gender specification based on the number of children and reports parametric evidence of bias against girls on consumption expenditure. However, that study did not address the robustness of that evidence to the lead total expenditure

[9] Collected in 2001, the data sets for this study includes households with completed fertility defined by mother age older than 50 years.

variable in presence of potential non-linearity in the adult goods Engel curves; nor did it consider the potential endogeneity of total expenditure and gender-affected fertility. The preset study addresses both issues.

3-Specification: semiparametric and parametric models

Much attention has been devoted to flexible or functional form free specification of (log.) of total expenditure as the lead explanatory variable in Engel curve analysis; known after Deaton and Muellbauer (1980) as the basis of the *almost ideal demand system* (*AIDS)*. The Working-Leser model with its logarithm of income specification (see Working, 1943 and Leser, 1963) has provided the main model for the estimation of child gender demographic effects using Engel curve analysis (see Deaton, 1997 for its application to a number of developing countries). As argued in section 2.1, one source of misspecification is the linearity of the log total expenditure term in that model, which suggests that non-linear extensions of the Working-Leser model with an additional quadratic term for expenditure (QAIDS) may provide greater functional form flexibility (see, for example, Dickens et al.,1993). However, evidence from specification-free methods suggests that the quadratically-extended Working-Leser models do not always adequately capture nonlinearities in Engel curve behavior. Pendakar (1999) examines the hypothesis that the quadratic Engel curves of demographically different households' budgets share the same curvature; the reported tests of quadratic Engel curves against their semiparametic specifications provide evidence for the superior performance of the latter. Blundell et al. (2007) extend that analysis by allowing for endogeneity of total expenditure and demonstrate the importance of nonparametric and semiparamertic specifications for isolating demographic effects in consumption behavior. The importance of a correct parametric specification of child gender in the Rothbarth model should be established only after the results are compared with semi-parametric specification of total expenditure as the principal variable in the Rothbarth-based Engel curves. While a robust approach to Engel curve estimation is valuable in offering functional-form free evidence on child gender bias, their significance in uncovering gender effects in consumption must be established in comparison to misspecification of gender demographics.

However, misspecification of functional form for demographic variables has received much less attention. If the source of misspecification is in the demographic variables, then a semiparametric approach may still fail to provide consistent estimation. Banks et al. (1997) provide good evidence in support of a flexible quadratic Working-Leser specification, but only if the model removes demographic variation by estimating Engel curves for single homogenous groups with total expenditure as the sole independent variable. Lyssiotou et al. (2008) demonstrate that, with greater functional form flexibility for demographic variables, there is no need for more flexible or nonparametric specifications of total expenditure in Engel curve estimation. In an study of food Engel's curve, Gozalo (1997) maintains that parametric misspecification of demographics are at least as importance of that of total expenditure; the paper tests for gender demographic specification and concludes that a more flexible parametric specification of demographics is required to obtain unbiased estimates, even with a nonparametric term for total expenditure. These studies suggest that parametric misspecification of demographic variables can be as important a source of misspecification as total expenditure; in this paper both are examined; the former with respect to child gender. We adopt a semiparametric strategy that combines a gender parametric demographic specification based on number of children and nonparametric functional form for total expenditure.

A frequently employed method among several available semiparametric models is the partially linear semiparametric estimator of Robinson (1988).[10] The Robinson semiparametric model is written as equation (1).

$$w_j = F(\ln x/n) + \beta.z + \varepsilon_j \qquad (1)$$

where w_j is the budget share of an adult good j, e.g., tobacco, x/n stands for per capita total expenditure, z is a vector of all other variables (e.g., demographics, regions, etc.), and F is the unknown function for x. In the first step, the conditional means of $E(w|\ln x)$, and each $E(z|\ln x/n)$ are estimated nonparamertically. In the second stage, $E(w|\ln x/n)$, and each $E(z|\ln x/n)$ are subtracted from both sides of the equation, which yields equation (2).

$$W - E(w_j|\ln x/n) = \{z - E(z|\ln x/n)\}.\beta + \varepsilon_j \qquad (2)$$

Equation (2) is the Robinson estimator, with the left-hand differenced-variable regressed on a vector of differenced z variables. The OLS estimation obtained when estimating equation (2) is unbiased, consistent, and asymptotically normal. The semiparametric specification for equation (2) with additional parametrically-specified demographic variables results in equation (3).

$$w_j = F[\ln(\frac{x_h}{n_h})] + \eta \ln n_h + \gamma_g \frac{n_g}{n_c} + \sum_{a=1}^{4} \delta_a \frac{n_a}{n_c} + \rho.\mu + \delta z + \varepsilon_j \qquad (3)$$

10 Yatchew and No (2001) offer a simpler method than Robinson (1988), as its endogneity test requires the nonparameteric component to be estimated only once. However, its differencing method results in higher standard errors than Robinson (1988).

where x_h is household total expenditure, n_h is household size, n_a/n_c indicates the proportion of children in different age groups a, over total number of children n_c. n_g / n_c is the proportion of girls in the household and constitutes the child gender index. We note that while $log\,(x_h / n_h)$ and $log(n_h)$ are defined with reference to n_h , both change only with number of children, not number of adults, since all observations consist of two adults plus $n_c \geq 1$. z is the vector of additional explanatory variables: regional dummies; variables specific to the analysis of child gender bias, such as work status of adults, dummies for the mother's age and her educational level; their interactive terms, etc.; ε_j is a random error term.

We add to (3) an additional variable μ on the right that controls for endogeneity of expenditure, or for that on number of children; unlike the other exogenous variables in (3) and (4), the observations on μ are estimated values obtained at a prior stage from the residuals of equations specified for each endogenous variable in (3) and (4). In this study we employ three versions of μ; two IV-based measures, one for $log\,(x_h / n_h)$; and one for $log(n_h)$; and a third obtained from the residuals of a binary logit model for probability of having more than one child conditional on the presence of at least one boy already in the household. We shall take up the specifications for the first stage function for each endogenous variable in (3) and (4) in the next two sections.

The coefficient of the child gender variable indicates the amount by which average expenditure share on an adult good is expected to change if a son is replaced by a daughter; holding constant the age structure of the children. More specifically, if adults are less willing to reduce their standard of living, as indicated by their consumption of adult goods, when there are more girls in the household (i.e., when n_g / n_c increases), then one would expect this proportion to have a positive effect on the budget share of each adult good (i.e., the higher the n_g / n_c ratio, relative to that for boys, the greater the expenditure on adult good *j*). Hence, the maintained and alternative hypotheses are: $Ho: \gamma_g \leq 0$ *vs*. $Ha: \gamma_g > 0$.

We test our hypothesis with the model specified in equation (3) against the semi-logarithmic Working-Leser flexible functional form augmented with a quadratic term for a curvature effect of the logarithm of per capita total expenditure. i.e., *QAIDS* specification.

$$w_j = \alpha_0 + \alpha_1 \ln\left(\frac{x_h}{n_h}\right) + \alpha_2 \left[\ln\left(\frac{x_h}{n_h}\right)\right]^2 + \eta \ln n_h + \gamma_g \frac{n_g}{n_c} + \sum_{a=1}^{4} \delta_a \frac{n_a}{n_c} + \delta z + \rho.\mu + \varepsilon_j \qquad (4)$$

Section 5.2 provides least square estimates obtained from the estimation of equations (3) and (4) for each *j*=4 adult good in this study. All expenditures are measured on annual bases, and the definition of total expenditure employed includes imputed rent but excludes housing investment. We note that equations (3) and (4) employ separate demographic proportions for child age and gender, combined into the same indices in earlier studies, with child gender defined here over the number of children. Five different child age groups are employed: under 2 years of age, 3-6 years, 7-9 years, and 10-12 years. We omit the 13-15 years cohort as the missing group to avoid collinearity. Similarly, there are two gender proportions of children, for boys and girls; the one included above is the proportion of girls. We include, in the vector of variables in *z*, controls for urban/rural, and for the latter, whether the household resides in a rural area under the safety net system run by the government and, if so, whether in addition the area has a free school feeding program.[11] The list of independent variables also includes controls for regions. For the 10 regions in the survey, 9 regional dummies are employed in equations (3) and (4), with Addis Ababa acting as the missing, reference region. We employ a specification test of equation (3) against equation (4) by employing the Hardle and Mammen procedure (1993) with a χ^2 test statistic based on the squared deviations between the fits of the semiparametric and parametric models and obtain critical values for the test from simulated values resulting from the application of wild bootstrap.[12] Failure to reject the null hypothesis suggests the polynomial employed (in the parametric model) is at least of the degree tested. Finally, it should be noted that the inclusion of non-purchase observations in the sample allows sharper results that capture both extensive and intensive margins of gender-related consumption of adult goods. For this reason, the relevant approach, as in previous applications, is to employ all positive and zero observations together, and it is the one adopted in this study. [13] Gong et al. (2005), for China, and Bhalotra and Attfield (1998),

---

11 Ethiopia's public works program known as the PSNP is the largest in Africa (see Gilligan et al., 2009), in addition in some of its areas school feeding programs are available. Both appear to significantly affect the Ethiopian consumption patterns, see tables 4 & 5 below. The list of rural areas in safety net regions, some with school feeding programs, was separately available. This was then matched with the survey regions/areas to construct safety net and school feeding dummies for the analysis.

12 Wu's (1986) wild bootstrap procedure obtains fitted values by resampling from the residuals while retaining the regressors at the sample means. Hardle and Mammen (1993, pp. 1933-1934) maintain that wild bootstrapping fitted values resolves the inconsistency resulting from simple bootstrapping procedures based on resampling from the original sample.

13 As two-step decision approach has had a mixed record of successful identification of the incidence of zero consumption for applications to expenditure data. For example, from estimates on educational expenditure in Kingdon (2005), it is unclear whether the evidence from the two-step method provides significant additional insight on gender bias beyond that offered by single-equation school enrolment estimation since none of the negative sign on girl effects on educational expenditure are statistically significant. The identification becomes a particularly acute problem with application to some key adult goods such as tobacco and alcohol, see Atkinson *el. at* (1984), because the zero consumption observations are a combination of non-consumption and measurement error zeros due to infrequency of purchase left out by the survey adopted cycle of

for Pakistan, are among a few studies that have examined child gender effects on Engel curves that employ nonparametric and semiparametic models similar to those employed in this paper, and both cover evidence on food as well as the Rothbarth Engel curves—tobacco and alcohol in the former and aggregate adult goods in the latter. Gong et al. (2005) report gender inequality in school attendance but find no matching evidence in terms of intrahousehold inequality that would provide an obvious explanation for that welfare outcome. In a somewhat similar semiparametric unpublished study (available on request), Koohi-Kamali and Liu (2014) report evidence of child gender bias by the Rothbarth model for China. Bhalotra and Attfield (1998) provide an interesting attempt for an alternative child gender specification based on birth order, but also report no evidence of gender bias effects on Engel curves.[14] The results in this paper therefore offer a means of evaluating the estimation and testing strategies adopted in this paper in terms of comparison with these earlier studies.

4-Endoneneity: total expenditure and demographics

The above model must consider the likely correlations between the error terms of the budget share equations with *a*) household total expenditure; *b*) household demographics, and *c*) the exclusion of the relevant explanatory variable from the budget share equations. Regarding *a*),the measurement error in total expenditure often arises either from infrequently purchased consumption goods, or from the survey sample recall periods longer than the typical cycle of household consumption. This issue assumes a larger significance in a predominantly rural developing economy since a great part of a household

---

consumption, see Deaton and Irish (1986).

14 It is notable that Pakistan usually has some of the worst indicators of gender welfare outcome.

consumption consists of home-produced products on own land without going through the market. In the absence of income data, as in the current study, wealth-related variables can act as alternative sources of instruments to control expenditure endogeneity. Assuming the separability of saving decisions over different time periods from consumption expenditure decisions within a given period allows treating true total expenditure as a function of wealth. Using wealth instruments also resolves the problems with the two sources of measurement error above since they are not subject to either type of error. However, wealth instruments must be robust with respect to under-reported, own-produced consumption; important typical rural wealth, land or cattle, may be sensitive to that source of measurement error; additional wealth instruments such as ownership of consumer durables provide the required robustness, particularly if the survey employed also covers urban households.

Turning to *b*), with son-preference, the number of children becomes a household choice variable, given the absence of pre-natal technology of child gender selection. There may be other potential sources of demographic endogeneity generated from the correlation between child age or birth spacing of children in the household and the error term of the budget share Engel curve equation of each adult good. Moreover, the specification of Engel curves must additionally consider the endogeneity of the number of children and the gender of children. Given the scarcity of instruments for demographic endogeneity available from a typical cross-section survey, consistent IV estimation is harder to achieve than it would be with expenditure endogeneity. There are however potential methods to address this problem. One solution for cross-sectional endogeneity demographic variables is to instrument just one of them; in the context of fertility and labor supply with cross-section data, Browning (1990, p. 1467) recommends that choice as a "potentially useful" solution, for evidence on its application see Cramer (1980). As a second solution, in the context of Engel curve estimation, we also discuss Gronau (1991) proposal for a non-IV approach to deal with number of children as a choice variable in estimating a budget share equation for adult goods.

The distances to public facilities and time spent to access these provide potential sources of instruments to control endogeneity. However, the estimation methods should distinguish between the instruments and the relevant excluded explanatory variables in the Engel curve budget share equations; and that would control for the third source of correlation with the error term in (3) and (4), namely that from *c*) above. In developing poor countries, the number of children is likely to be correlated with the time to the nearest primary school and health center facilities but such variables are also likely to leave the budget shares of adult goods unaffected.[15] A related set of instruments, in terms of time expenditure, are those that reflect the opportunity cost of education. We further employ time to and distance from the nearest source of drinking water collection, time spent on fire-wood collection for heat and cooking as additional instruments for demographic endogeneity; these are likely to prove useful since much of the responsibility for wood and water collection rests on the shoulders of female children, often an important incentive for withdrawing girls from school. Finally, the IV application to control for potential endogeneity of total expenditure and household size must allow for functional form

[15] Attanasio and Lechene (2010) employ similar instruments to deal with endogeneity of school attendance as a qualifying condition for a safety net program in rural Mexico.

flexibility and presence of non-linearities in the budget share Engel curves by necessary modification of the standard linear IV application.

We employ control function approach of Blundell *et. al*. (1998) , see also Holy and Sargan (1982), for this purpose, using the residuals from the first stage least squares IV estimation as an additional variable in the budget share equations in the second stage to correct for inconsistent estimates with non-linear Engel curves, though the method is equivalent to IV in the linear case. Moreover, as its by-product, this approach also provides a test of over-identification endogeneity, given the set of employed instruments: a residual significantly different from zero would suggest a rejection of instrumental exogeneity with respect to the budget share equations; hence providing evidence on whether time-based instruments affect expenditure decisions on adult goods in their own right. Attanasio and Lechene (2014) applied the technique to a resource share of food of each member in a collective model of the household, while Gong el. at. (2005) employed the approach with a household-level Engel curve for adult goods.

We specify the following two instrumental equations in order to obtain the control function residuals employed in (3) and 4).

$$\ln\left(\frac{x_h}{n_h}\right) = \alpha_0 + \sum_{l=1}^{l} \alpha_l w_l + \sum_{j=1}^{j} \beta_j z_j + \mu_x \qquad (5)$$

where w_l stands for the *l*th wealth instrument for expenditure; z_j stands for each one of *j* exogenous variable in (3) and (4) acting as its own instrument, and $\mu_x(0,\sigma_x^2)$ is a random error term for the expenditure instrumental eq. (5).

$$\ln(n_h) = \alpha_0 + \sum_{d=1}^{D} \alpha_d\, dist_d + \sum_{t=1}^{T} \beta_t\, timex_t + \sum_{m=1}^{M} \gamma_m z_m + \mu_c \qquad (6)$$

where $dist_d$ stands for the *d*th distance instrument for $\ln(n_h)$, $timex_t$ stands for the *t*th time expenditure instrument; z_m stands for each one of *m* exogenous variables in (3) and (4) acting as its own instrument, and $\mu_c(0,\sigma_c^2)$ is a random error term for household size instrumental eq. (6).

Since the performance of IV for demographics with a cross-section data is not as effective as that for total expenditure; application of an alternative non-IV solution is useful in demonstrating robustness. We also test an alternative to instrumenting demographics for *b*) above. Gronau (991) suggested to deal with the possible endogeneity of number of children by adding the residual from an estimation of a logit model estimation of the probability of having children, defined as a binary variable, to the budget share equation of adult cloths as a new explanatory variable to control adults' "taste" for children; controlling separately for the impact of an increase in number of children on sources available to adults by adding an interactive term of the residual and consumption expenditure. We follow a similar strategy defining two different binary variables to examine the impact of son-preference on the number of children. For logit depend variable we define a dummy *K* equal to 1 if number of children > 1, 0 otherwise. The principal conditioning variable would be the presence of a boy, defined as a dummy equal to 1 if the child is the *oldest* boy in the family, zero otherwise.[16] The specification of the logit function for *K* is as follows:

$$\ln\left(\frac{K}{1-K}\right) = \alpha_0 + \alpha_{1st} boy_{1st} + \sum_{m=1}^{M} \beta_m x_m + \mu_k \qquad (7)$$

[16] We also examined the results with a simpler dummy equal to 1 if the family has at least one boy. The logit residuals from either one of the main conditional binary variable for a boy presence turn out to have very similar impact on the estimates in the budget share equation. We decided to employ the estimates based on the oldest boy dummy as the more sensible of the two.

Where boy_{1st} is the oldest boy present in the household, x_m is the mth additional control, e.g. mother's age or education. The logit dependent variable estimates the probability of having more than one child $P(K=1)$ from log of the odds, $log\,(\frac{\hat{K}}{1-\hat{K}})$, conditional on the dummy for boy_{1st} ; plus additional controls discussed in the next section. The residual from that regression, namely $\mu_k = (K - \hat{K})$, is a measure of deviation from that probability; the greater the deviation, the more important the unobservable adults' "taste" effects for the number of children presence. We include μ_k as an additional explanatory variable in (3) and (4) in order to control for potential endogeneity effects of number of children on adults consumption. Moreover, with log of per capita total expenditure constant, a change in the number of children also affects with the per capita total expenditure available per child, we also include in (3) and (4) an interactive term of μ_k with consumption expenditure to control for the impact of number of children endogeneity on the household total expenditure. This seems useful in separating a change in the adults' share of household resources due to a change in number of children from that of a change in the gender of children, see also Gronau (991) and discussion in section 2.2 above on the effects of more girls in larger households on household per capita income, given equal gender treatment. A statistically significant parameter for this variable (ρ_k) in the share equations based on (3) and (4) would indicate number of children endogeneity in this approach. The advantage of this method is that its implementation does not require the existence of effective instruments for demographics that are usually difficult to come by.

To sum up, we examine endogeneity in (3) and (4) with three measures of μ: μ_X obtained with wealth instruments from (5) for expenditure, μ_C obtained with time and distance instruments from (6) for number of children, and μ_K obtained from a logit function for the probability of having more than one child on conditional on the oldest boy in the household from (7), also for number of children.

4-Data and sample truncations

The data for this study comes from the 2004-2005 Household Income and Consumption Expenditure Survey (HICES) conducted by the Ethiopian Government's Statistical Authority. The data set consists of over 21,299 households and the survey covers the entire country. The data set contains extensive expenditure information but *no* income data. However, the samples employed are confined to nuclear household observations and are smaller than the full sample. This was to avoid misleading effects associated with extended family observations that are commonly included in most studies on Engel curve estimation of child gender effects; for instance, all studies cited in Deaton (1997) appear to employ samples inclusive of extended households.

The typical developing country to which the model has been applied is one with a very large percentage of multi-adult extended households. The application of the model would lead to distorted results when applied to surveys from developing countries if the observations on adults' consumption are treated as though they reflect solely the allocational decision of the parents regarding the gender of their children. This would violate the implicit assumption underlying the Rothbarth model, which presupposes that expenditure patterns reflect the preferences of the parents as the sole decision-makers (see Deaton and Muellbauer,1986). In other words, the Rothbarth model is applicable to internal allocation within a *nuclear* family.[17] An indiscriminate application of the model to all types of households, as in all previous attempts,[18] is likely to distort the evidence of child gender bias in consumption patterns.

This study employs only samples of nuclear household units. The definition of a nuclear household we employ for sample truncation consists of the household head and the spouse and at least one child under 15 years of age. Thus, households that contain children above 15 years of age are excluded. Table 1 shows the percentages for different types of households. Nuclear households with at least one child, account for 52% of the total. However, about a third of such nuclear households, accounting for 17% of the entire survey households, are those that contain non-biological children of the head or spouse. Some studies suggest that parents discriminate among children by their degree of "genetic relatedness," resulting in poorer treatment for non-biological children than for biological ones, see Case et al. (2000). Koohi-Kamali (2017) considers Ethiopian (mainly HIV) orphans and provides evidence that girls face more severe discrimination than boys; even biological girls are found to fare worse than orphan boys. However, unlike the effect of biological children, taking orphans into a household is a deliberate decision; and that introduce into the model an important endogeneity problem, that is best set aside in a study

[17] The extended family is a major household type in most developing countries; and in many likely to constitute the principal unit of consumption. For example, Lanjouw and Ravallion (1993, table 1) for rural Pakistan classify demographic groups by the number of adults. If we assume there are no extended families in groups with up to three adults, and households with four or more are exclusively of non-nuclear types, then the extended families account for 56% of their sample.

[18] Gong et al. (2005) is an exception; in this regard; the collective models discussed above have adopted a more careful sample truncation rules.

that is concerned mainly with the effectiveness of the Rothbarth internal inequality model in the treatment of children by gender. As it is easier to work with a standard biological parents-children setup, we have confined the sample employed in this study to those with biological children of the head and spouse. The resulting sample includes 7890 households.

Four categories of expenditure appear reasonable candidates for adult goods in Ethiopia: tobacco, coffee, adults' clothes, and adult personal services and personal effects, "personal services" for short. (The number of observations of alcohol was very limited.) The latter provides consistently the strongest evidence for child gender bias in Ethiopia. The items in this group are made up of shaving-related goods and services, in addition to perfumes, handbags and wallets, walking sticks, and wigs. Table 2 shows the incidence of purchase/non-purchase for each group in the survey, ranging from just under a quarter to well over a third of the households. As such, the percentages of non-purchase are not unlike those expected of some common adult goods; see Subramanian and Deaton (1991) for India. Finally, Table 3 gives the mean and the standard deviation of the main variables of equations (3) and (4) for the main, all-country (covering both urban and rural areas) sample of nuclear households consisting solely of a household head, their spouse, and their biological children under 15 years of age.

We instrument log of household total expenditure by land ownership, and two items of consumer durables: ownership of a cooking stove (gas/electric) and of a radio, with expected positive sign in the first stage regressions. We instrument log of household size by time to the nearest primary school, and health center, and sources of firewood and water collection. Moreover, availability of transport generates further variation across households in time spent on frequent trips to the same school distance; hence we also include primary school distance as an additional variable. Since the sample employed consist of two-parent nuclear households, household size varies only with the number of children. We expect positive sign for expenditure instruments. It is harder to suggest a priori sign for the number of children instruments. We expect a positive sign for drinking water distance since more children make collecting water collection in longer distance more manageable for the household. We also expect a negative correlation between the distance to school and the number of children as a longer distance makes school attendance harder, especially if children have to attend different schools. These instruments are assumed exogenous in the instrumental equations, and to resolve the violation of this assumption due to potential measurement error of the instrumental variables, we employ all instruments at their average values in the smallest administrative unit (wereda) defined for rural and urban areas; for the sample of 7865 households employed in this study, there are 1562 such units.

Finally, the logit residual term μ_k, for the probability of having more than one child conditionally on presence of the oldest boy obtained from (7), is added to (3) and (4) as an alternative to μ_c for household size endogeneity. We also included the following additional controls, separately for the head and spouse, for the vector x_m in (7): age and age-squared, employment and marital status; plus log of per capita household expenditure, linear and quadratic, dummies for rural-urban, school-feeding & public work schemes and nine regional dummies. We examined the impact of μ_k, on the budget share equations based on (3) and (4) once by including μ_k, alone, and once by including both μ_k number of children and μ_x for expenditure together. The results on the estimates in (3) and (4) are similar, see Appendix B, table 2c. To save space, below we report those with both μ_k and μ_x included in (3) and (4). The application of (7) also required a decision on the sample twin births. As the sample reports age in full years, we lack the information to distinguish between twin children and children borne in the same years; we treat all “twins” (just over 300 observations in the sample) as single-births for the purpose of defining a binary boy presence variable.

5-Results

We examine the results in two parts. The nonparametric section provides some preliminary comparisons between the Engel curves of boys- and girls-dominated households. This is presented in section 5.1. We also present the estimates and tests for nonparametric specification of log of total expenditure and parametric specification of the number of children. These results are presented in section 5.2 (separately with total expenditure and demographic endogeneity controls in section 5.2.1 and 5.2.2 respectively).

5.1-*Nonparametric results*

Some primary evidence on higher budget shares of adult goods for households with more girls than boys, based on specification-free child gender effects on consumption patterns, is presented below. These results are obtained from comparisons of nonparametric regression of budget shares of 1) adult goods and services, 2) adult clothing, 3) tobacco, and 4) coffee on the logarithm of per capita total expenditure separately for households with more female than male children and those with more male than female children. The regression plots are based on the Gaussian kernel function of order 2, and the optimal bandwidth was determined by integrated squared error procedure (see Silverman, 1986). In order to contain the effect of observations at the tails of the function, the sample is trimmed at 1% for low densities.

A pairwise comparison of plots displays a general pattern for the girl-dominant curves to be *higher* than the boy-dominant curves, consistent with the predicted direction of female child discrimination. This pattern is more notable in plots of adult goods and services; moreover, the girls' group curve is also higher at the top end of the expenditure scale for adult clothing and at the bottom end for coffee and tobacco. However, it is not obvious whether the differences between the nonparametric equations are statistically significant or not. In particular, without controlling for endogeneity the interpretation of these nonparametric regression results can be misleading. The next section formalizes these nonparametric differences by constructing a semiparametric estimator discussed above. However, additional covariants such as child age, parents' age and education, etc. must be controlled in order to obtain more clear measures of the effect of child gender on the household consumption patterns. Semiparametric and parametric evidence with an extensive list of additional controls are presented next.

5.2- *Semiparametric and parametric*

Tables 4, 5 and 6 present the semiparametric estimates for equation (3) and parametric estimates for equation (4) by *QAIDS* specification, for gender effects on consumption; first with control for (log) total expenditure endogeneity (table 4), then with functional form approach to control for (log) of household size, that is the number of children, endogeneity (table 5), and finally with both logit residual control for number of children and functional form to control approach for consumption expenditure. The first column under each category of adult goods presents the semiparametric estimates of equation (3), and the second column presents the parametric estimates of equation (4). Here we present the main variables of interest; Appendix A shows the estimates inclusive of the key additional controls; Appendix B presents the first stage estimates for the instruments employed and those by the logit function.

5.2.1-Expenditure endogeneity

The main evidence on child gender bias appears in the first (highlighted) row of the table 4 with control for expenditure endogeneity. The semi-parametric coefficients estimates (left column) have the expected positive sign for bias against girls for all the four adult goods estimates of this study. All four coefficients are statistically significant, especially for the category of adult personal items and services. The household size effect, usually interpreted as a measure of economies of scale in consumption, is positive but statistically significant only for adult clothing and coffee. The endogeneity coefficient estimate, ρ_x, is statistically insignificant for all four adult goods, suggesting the log of per capita total expenditure in equations (3) and (4) is exogenous; Appendix A shows that the statistically significant instruments trend to have the expected positive sign, expect for adult cloths. The corresponding parametric results (right column) provides the corresponding parametric estimates. Once again, the child gender effects have the expected positive sign and are statistically significant for all four adult goods. However, the clothing equation is log-linear.[19] The coefficient on the log of household size is positive and statistically insignificant for adult goods but tobacco (parametric equation).[20] We test the

19 The estimation of (4) by excluding the quadratic term for adult clothing results in a linear with very similar parametric estimates for the remaining variables in (4).

20 Parametric evidence is also tested for sensitivity to the precise age definition of a child, due to possible effects of child labor on adult goods consumption, with a child age restricted to 12 or less.

semiparametric model against the quadratic equation (4) for specification in table 4. The results indicate the rejection of the parametric model in both linear and quadratic versions for coffee and for tobacco, and for personal items for the linear but not for the quadratic version.

5.2.1-household size endogeneity by control function approach

Table 5 presents the semiparametric estimates (left column) for equation (3) and *QAIDS* equation (4) (right column), with control for (log) of household size (number of children) by control function approach. The main evidence on child gender bias appears in the first (highlighted) row of the table. These coefficients have the expected positive sign for bias against girls for all the four adult goods estimates of this study. All four coefficients are statistically significant, especially for the category of adult personal items and services. The household size effects are positive but statistically insignificant for all share equations; moreover, the endogeneity coefficient estimate, ρ_c, is statistically insignificant for all four adult goods, suggesting the log of per capita total expenditure in equation (3) is exogenous; Appendix B, table 2b shows that the school and water location distances have the expected negative and positive signs. The right column of the table, under each adult goods category, provides the corresponding parametric estimates. Once again, the first row presents child gender effects that have the expected positive sign and are statistically significant for all four adult goods. The coefficient on the log of household size is positive and statistically insignificant for all but tobacco; however, endogeneity term

---

Moreover, Ethiopia has a very large percentage of female-headed households, see Desta et al. (2006), and some studies suggest that a mother's control over household budget improves child welfare. Parametric estimation is also repeated with a smaller sample of female-headed households. The results on child gender proved to be insensitive to these changes (see Koohi-Kamali, 2008).

for household size also remain all insignificant, suggesting size is exogenous in equation (4). We also test the semiparametric model against the linear version of equation (4) for specification; the results indicate the rejection of the parametric model in both linear and quadratic versions for coffee and the parametric linear model for personal items and cloths.

5.2.2-household size endogeneity by logit residual control

Table 6 presents the semiparametric estimates (left column) for equation (3) and the *QAIDS* estimates (right column) with control for both (log) of household size (number of children) by the logit residual and log of per capita total expenditure. The main evidence on child gender bias appears in the first (highlighted) row of the table. These coefficients have the expected positive sign for bias against girls for all the four adult goods estimates of this study. All four coefficients are statistically significant, especially for the category of adult personal items and services. The household size effects are positive and statistically significant in adult cloths and coffee share equations. Moreover, neither of the two endogeneity coefficient estimate, ρ_x for consumption expenditure and ρ_k for the number of children, are statistically significant for all four adult goods, suggesting the log of per capita total expenditure in equations (3)-(4), and number of children are exogenous in equations (3)-(4). Table 3a in Appendix A presents a fuller list of parameter estimates inclusive of an interactive term between ρ_k and consumption expenditure; we note that the interactive remain insignificant in all four share equation; separate estimates with endogeneity control for number of children alone by logit residual, not reported here and available from the author, also produced very similar estimates.

To sum up, endogeneity controls are insignificant throughout with both (log) total expenditure and the number of children; the specification test are particularly important for coffee, and for tobacco (when expenditure is instrumented), suggesting the *QAIDS* functional form is not sufficiently flexible to capture the non-linearity in the Engel curves of these goods. However, the quadratic specification rejection is more limited with household size endogeneity than with expenditure endogeneity. The linear versions of the model are more extensively rejected with either type of endogeneity controls. However, linearity is not rejected for cloths in tables 4 and 6, and for tobacco in table 5, see Gong et al. (2005, p. 10) for similar outcomes for alcohol and tobacco. The child gender effects on adult consumption, however, remain robust regardless of the sources of endogeneity and regardless of the Engel curve specification for its lead variable; in particular the children gender effects remain similar with number of children endogeneity by control function or or logit residual approaches. This outcome lends support to the proposed specification of child gender based on the number of children; although instrumenting demographics is not as effective as instrumenting total expenditure., the alternative of employing the deviation of actual from probable number of children goes some way to back up the *IV* evidence on number of children exogeneity. Finally, we report that the correlation coefficient between the number of children and gender proportion in the sample employed turns out to be close to zero and insignificant: the Pearson correlation coefficient = -0.0107 / sig. prob. = 0.3426.

6-Conclusion

This study re-examines the Rothbarth model of inferring child gender discrimination from household consumption patterns with the main objective of demonstrating its effectiveness in view of the failure in most previous applications to uncover evidence of child gender bias. We maintain as a working hypothesis that child gender specification should be based on the number of household children rather than their age. We employed parametric and semiparametric Engel curves for the Rothbarth model of intrahousehold child gender disparity with nonparametric and non-linear parametric specifications for log of total expenditure, and parametric specification for gender based on the number of children rather than their age. The paper attempted to demonstrate the effectiveness of the proposed approach with an application to a 2004-2005 Ethiopian household expenditure survey based on a sample of strictly nuclear households with biological children.

The results suggest evidence for the presence of effects of bias against girls on consumption, to various degrees, by all four adult goods available to this study. The group of adult personal goods and services provides the strongest and the most consistent evidence of discrimination against girls in Ethiopia. In addition, evidence from other adult goods in this study is notable for suggesting similar effects, though these are typically not as strong. We also tested the semiparametric against the parametric models employing flexible functional forms, and the outcome in general supports the superiority of the semiparametric results: the *QAIDS* parametric model is rejected for adult goods and services and for coffee. However, once we control for number of children endogeneity by a logit residual, the *QAIDS* specification is rejected only for the budget share of coffee. Furthermore, the semiparametric estimates also include control for potential endogeneity of logarithm of per capita total expenditure and number of children, the evidence of gender bias remains largely unaffected by this control. Hence, the evidence for child gender on consumption remains robust to changing a) the specification for total expenditure and b) to different methods of endogeneity control, that is this study suggests that parametric misspecification for child gender demographic variable remains important for identification of child gender effects on consumption with or without free functional form for total expenditure and after controls for endogeneity. The results presented suggest public policies should be designed for and be targeted directly on female children rather than the households they are a member of.

Table 1- Percentage distribution of types of household in the Ethiopian 20045-2005 HICES	
Extended households	34.7% (7401)
Nuclear/biological children of head-spouse	34.3% (7298)
Nuclear/non-biological children of head-sp	17.6% (3751)
Nuclear/other types (one person, etc.)	13.4% (2849)
All nuclear types	65.3% (13898)
All household types	100% (21299)

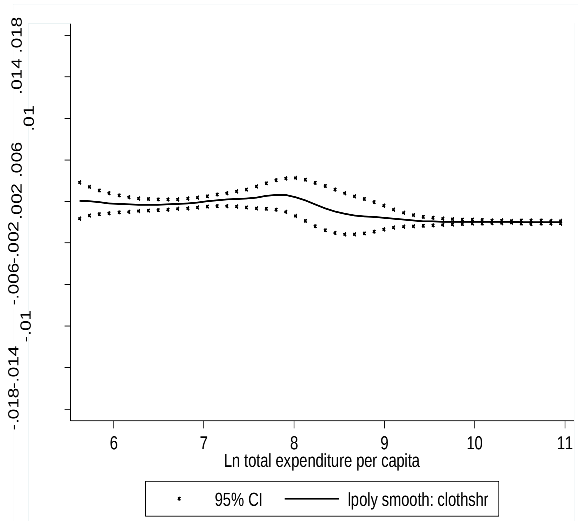
Table 2- Survey incidence of purchase/non-purchase of selected adult goods	
Adult Personal effects and services [a.]	24.2% (7606)
Adult Clothes	35.2% (5656)
Coffee	26.2% (5222)
Tobacco	36.5% (7872)

[a] Mostly shaving-related items and services, perfume, handbags, wallets, sticks, wigs, etc.

Table 3a. Mean, st. dev., and percentiles of budget shares & log of total exp.—all country					
Variables	Mean	Std. Dev	25th percentile	Median	75th percentile
Personl serv	0.00390	0.01110	0.00000	0.00000	0.00196
Adu.Clothes	0.00226	0.01680	0.00000	0.00000	0.00028
Coffee	0.01278	0.04574	0.00000	0.00000	0.00009
Tobacco	0.00594	0.02044	0.00000	0.00000	0.00312
Total expend	7.20124	0.60212	6.77310	7.15486	7.54961
Total exp^2	52.22039	8.95880	45.87490	51.19196	56.99667

Table 3b-Mean and st. dev. of the principle variables—the main all-country sample		
Variables:	Mean	Stand. Deviation
Household size	1.01038	0.61086
Proportion of children 0-2	0.22832	0.31236
Proportion of children 3-6	0.24911	0.27834
Proportion of children 7-9	0.14938	0.20215
Proportion of children 10-12	0.10646	0.17028
Proportion of girls	0.47541	0.33653
Urban/rural dummy	0.46641	0.49890
Safety net dummy	0.39899	0.48972
School feed dummy	0.17592	0.38078
Regional dummies:		
1-Tigray	0.06401	0.24478
2-Affar	0.04094	0.19816
3-Amhara	0.16781	0.37372
4-Oromiya	0.23726	0.42543
5-Somali	0.05957	0.23670
6-Benishangul	0.04715	0.21197
7-SNNP	0.16477	0.37099
8-Harari	0.03283	0.17819
9-Dire Dawa	0.03270	0.17786
Sample size	7890 households	

Plots of nonparametric regressions of the adult goods budget shares on log. of per capita total expenditure based on the local polynomial smoothing of Fan (1992) using the Gaussian kernel with 1% trimming at low densities. The bandwidth is obtained by the minimization of the conditional weighted mean integrated squared error. The 95% pairwise confidence intervals are constructed from the square root of the conditional variance of the local polynomial at each grid point (Fan and Gijbels, 1996, pp.116-118).



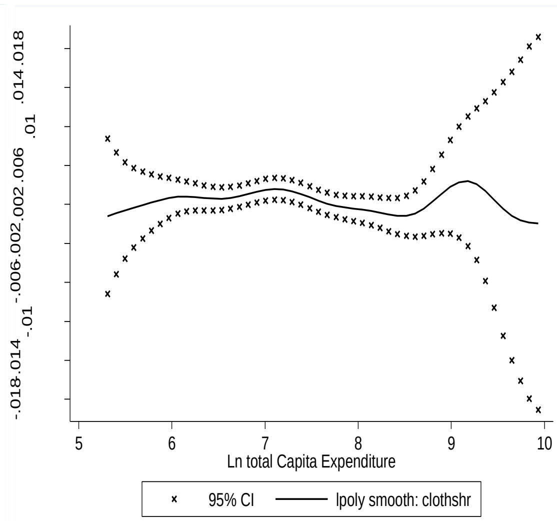
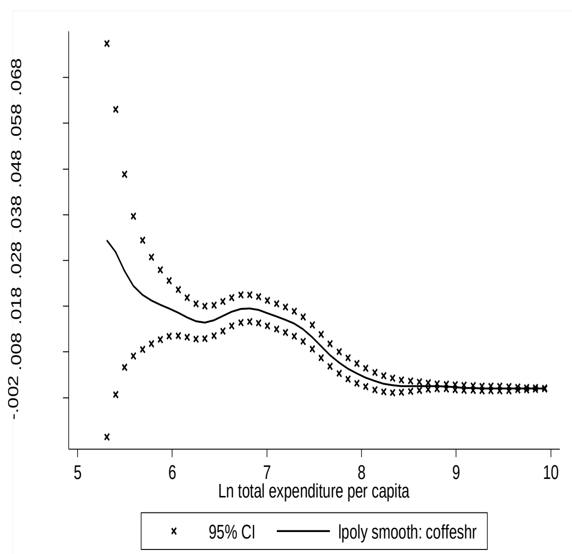


Fig. 1a Budget share*: Adult Clothing*, 95% CI
Families with the number of boys > number of girls

Fig.1b Budget Share: *Adult Clothing*, 95%CI
Families with the number of girls > number of boys .



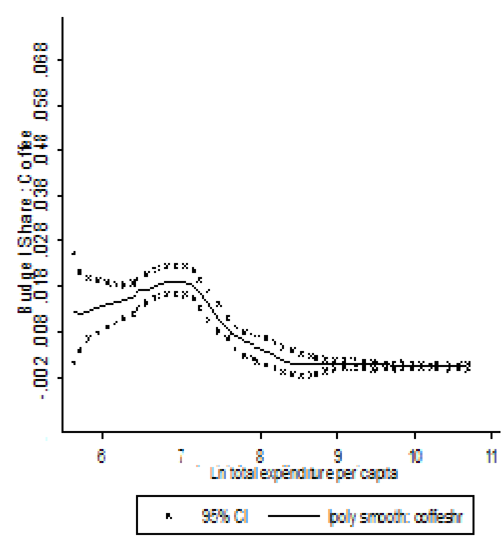
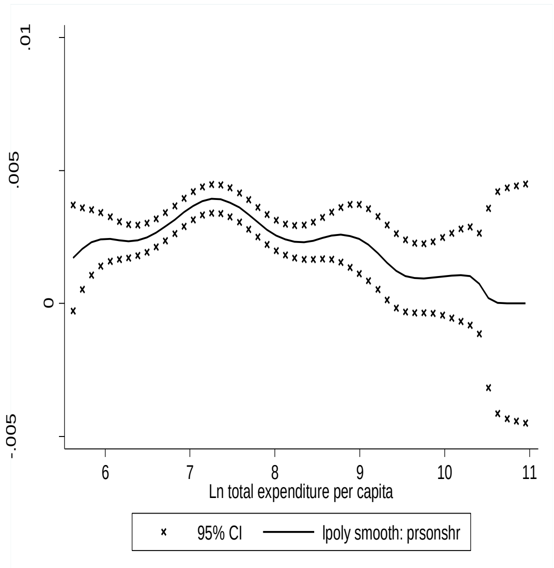


Fig. 2a Budget share*: Coffee*, 95% CI
Families with the number of boys > number of girls

Fig.2b Budget Share: *Coffee*, 95% CI
Families with the number of girls > number of boys



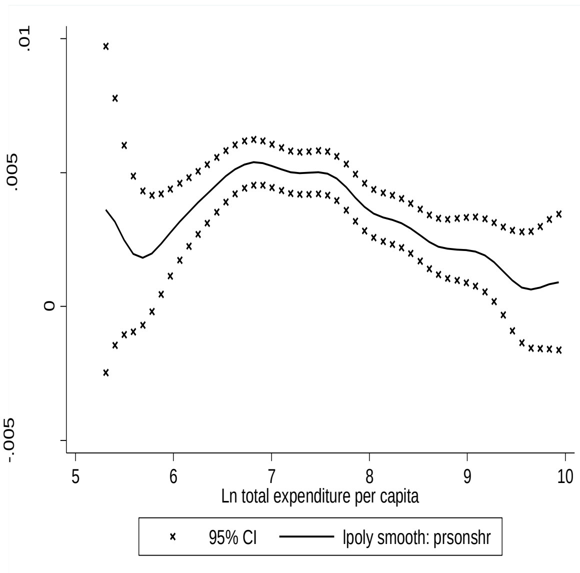
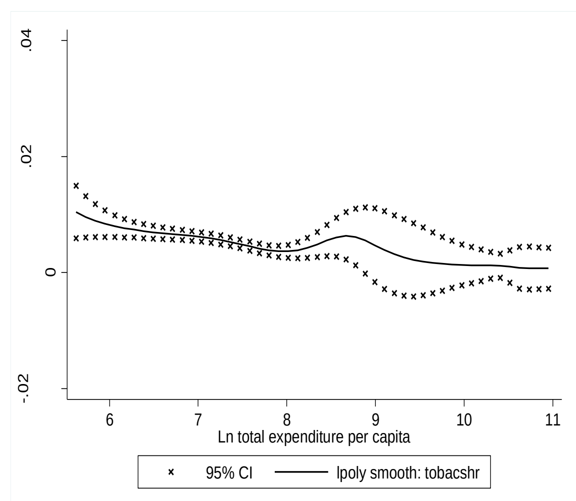


Fig. 3a Budget share*: Personal Services*, 95% CI

Families with the number of boys > number of girls

Fig.3b Budget Share: *Personal Services*, 95% CI

Families with the number of girls > number of boys



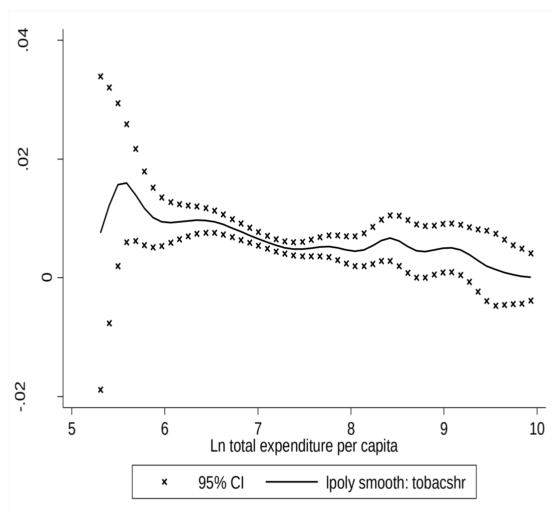


Fig. 4a Budget share*: Tobacco*, 95% CI

Families with the number of boys > number of girls

Fig.4b Budget Share: *Tobacco*, 95% CI

Families with the number of girls > number of boys

Table 4 Principal control variables for the budget shares of adult goods Engel curves (*t*-ratios in brackets); **total expenditure instrumented for endogeneity**

Dep. Var	Personal Services		Adult clothing		Tobacco		Coffee	
Variable	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet
(girls/n)	**0.0025****	**0.0024****	**0.0020***	**0.0021***	**0.0018****	**0.0018****	**0.0041****	**0.0040****
	(5.25)	**(5.18)**	**(2.03)**	**(2.05)**	**(2.37)**	**(2.80)**	**(2.38)**	**(2.36)**
ln n	0.0004	0.0004	0.0038	0.0038	0.0013	0.0013	0.0064	0.0064
	(0.39)	(0.39)	(1.80)	(1.83)	(0.58)	(1.01)	(1.87)	(1.87)
endogeneity	0.0014	0.0014	-0.0066	-0.0069	-0.0030	-0.0031	-0.0085	-0.0086
control ρ_x	(0.50)	(0.49)	(-1.26)	(-1.30)	(-0.48)	(-0.75)	(-0.72)	(-0.74)
Specific.	*Against semiparametric models by Hardle-Mammen procedure*							
Tests	(p-value in parenthesis)							
Linear spec.	7.6577 (0.0000)*		0.9132 (0.3900)		2.1196 (0.0433)*		3.3076 (0.0067)*	
Quad. Spec.	1.2440 (0.1567)		0.9190 (0.3600)		0.8984 (0.3133)*		2.5985 (0.0167)*	

Notes: Number of observations (after 1% trimmed at low densities) = 7837 consisting of nuclear households with one or more (biological only) children. Standard errors of all equations are controlled for clustering effect at the lowest administrative (wereda) units. Semiparametric estimates are obtained from Robinson (1988) $\sqrt{N}$-consistent partial linear procedure. Hardle-Mammen specification tests are for the linear and quadratic parametric models against the semiparametric model; ***** indicates rejection for 5% χ^2 critical values, generated by *300* replications of the wild bootstrap. Endogeneity parameter residuals ρ are obtained parametrically from *IV* applications with land as the principal instrument for total expenditure but also include ownership of a cooking stove and a radio; first stage estimates and *t*-ratios for land are 0.111 (4.60) adult goods and services, -0.0048 (1.74) adult clothing, 0.1024 (4.10) tobacco, and 0.1000 (4.10) coffee. Gender effect ** and * indicate significance at 1% and 5% respectively for a one-sided test of gender effects.

Table 5 Main control variables for the budget shares of four principal adult goods (*t*-ratios in brackets) **the number of children instrumented for endogeneity**

Dep. Var	Personal Services		Adult clothing		Tobacco		Coffee	
Variable	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet
(girls/n)	**0.0025****	**0.0025****	**0.0022****	**0.0022****	**0.0014***	**0.0015****	**0.0278***	**0.0277***
	(4.09)	**(4.04)**	**(2.42)**	**(2.40)**	**(2.05)**	**(2.44)**	**(1.66)**	**(1.64)**
ln n	0.0008	0.0006	0.0176	0.0168	0.0051	0.0039	-0.1102	-0.1098
	(0.11)	(0.09)	(1.39)	(1.30)	(0.57)	(0.27)	(-1.47)	(-1.45)
endogeneity	0.00003	0.0002	-0.0165	-0.0158	-0.0050	-0.0037	0.1140	0.1137
control ρ_c	(0.00)	(0.03)	(-1.28)	(-1.20)	(-0.55)	(-0.25)	(1.51)	(1.49)
Specific.	*Against semiparametric models by Hardle-Mammen procedure*							
Tests	(p-value in parenthesis)							
linear spec	7.6293 (0.0000)*		2.7279 (0.0000)*		1.4299 (0.1300)		1.8447 (0.0533)*	
quad. Spec	1.2415 (0.1600)		0.7936 (0.4300)		0.6434 (0.5133)		2.2963 (0.0333)*	

Notes: Semi-parametric estimation and tests as in notes to table 4. Endogeneity parameter residuals ρ_c are obtained parametrically from *IV* applications using the average time (per wereda) to the nearest primary school, health center, and sources of drinking water, fire-wood collection; and average distance to the nearest primary school and source of drinking water. Principal instruments (abs. t-ratio in bracket) are school distance: personal items 0.0068 (2.16), cloths -0.0048 (1.74), tobacco -0.0074 (2.47) and coffee-0.0008 (0.35); distance to water: personal items 0.0038 (3.42), cloths 0.0025 (2.42), tobacco 0.0025 (2.26), and 0.0012 (1.34); average time for wood collection: personal items -0.0022 (2.96), cloths -0.0013 (1.59), tobacco -0.0021 and coffee -0.0012 (2.05). Gender effect ** and * indicate significance at 1% and 5% respectively for a one-side test of gender effects.

Table 6 Principal control variables for the budget shares of adult goods Engel curves (*t*-ratios in brackets); **number of children endogeneity control by logit residual**

Dep. Var	Personal Services		Adult clothing		Tobacco		Coffee	
Variable	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet
(girls/n)	**0.0025****	**0.0025****	**0.0021***	**0.0021***	**0.0018****	**0.0018****	**0.0041****	**0.0037****
	(5.27)	**(5.69)**	**(2.06)**	**(2.08)**	**(2.36)**	**(2.83)**	**(2.38)**	**(2.18)**
ln *n*	0.0005	0.0005	0.0038	0.0040	0.0013	0.0013	0.0085	0.0083
	(0.57)	(0.55)	(1.84)	(1.86)	(0.58)	(1.05)	(2.36)	(2.34)
endogeneity	0.0014	0.0014	-0.0066	-0.0069	-0.0029	-0.0031	-0.0079	-0.0081
control ρ_x	(0.49)	(0.48)	(-1.26)	(-1.30)	(-0.46)	(-0.73)	(-0.67)	(-0.69)
endogeneity	0.0009	0.0007	-0.0010	-0.0011	0.0017	0.0019	-0.0039	-0.0034
control ρ_k	(0.70)	(0.60)	(-0.53)	(-0.69)	(0.44)	(0.65)	(-0.81)	(-0.75)
Specific.	*Against semiparametric models by Hardle-Mammen procedure*							
Tests	(p-value in parenthesis)							
Linear spec.	8.0881 (0.0000)*		0.9313 (0.3900)		2.1272 (0.0367)*		3.5014 (0.0067)*	
Quad. spec.	1.3897 (0.1133)		0.9352 (0.3533)		0.9102 (0.3200)		2.7960 (0.0133)*	

Notes: Semi-parametric estimation and tests as in notes to table 4. Endogeneity of number of children ρ_k are obtained parametrically from the residuals of a logit equation with dependent variable as *log* ($\frac{\hat{K}}{1-\hat{K}}$), estimating probability for a binary increase in the number of children beyond the oldest boy, in the household, $\mu_k = K - \hat{K}$; including its interactive term with expenditure to control effect of a rise in the number of children PC total expenditure. Gender effect ** and * indicate significance at 1% and 5% respectively for a one-sided test of gender effects.

Bibliography

Ahmand A, Morduch J. 1993. Identifying sex bias in the allocation of household resources: evidence from linked household surveys from Bangladesh. *Harvard Institute of Economic Research Discussion Paper* No. 1636.

Attanasio,, O, Lechene V. 2010. Conditional Cash Transfers, Women and the Demand for Food. *Institute of Fiscal Studies Working Paper No. 10/17*.

Attanasio,, O, Lechene V. 2014. Efficient Reponses to Targeted Cash Transfers. *Journal of Political Economy* **122(1):** 178-222.

Banks B, Blundell R, Lewbel A. 1997. Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics* **LXXIX**: 527-539.

Becker G. 1981. *A Treatise on the Family*. Harvard University Press: Cambridge, MA.

Becker G, Tomes N. 1976. Child endowments and quantity and quality of children. *Journal of Political Economy* **84(4, pt.2)**: S143-S162.

Bhalotra S, Attfield C. 1998. Intrahousehold resource allocation in rural Pakistan: a semiparametric analysis. *Journal of Applied Econometrics* **13**: 463-480.

Blackorby C, Donaldson D. 1993. Adult-equivalence scales and the economic implementation of interpersonal comparison of well-being. *Social Choice and Welfare* **10**: 335-361.

Blundell R, Chen X, Kristensen D. 2007. Semi-nonparametric IV estimation of shape-invariant Engel curves. *Econometrica* **75(6):** 1613-1669.

Blundell R, Duncan A, Pendakar K. 1998. Semiparametric estimation and consumer demand. *Journal of Applied Economterics* **13(5**): 435-461.

Bradbury B. 1994. Household semi-public goods and the estimation of consumer equivalence scales. Paper presented at CEPR/ERC.

Browning, M. 1992. Children and Household Economic Behavior. *Journal of Economic Literature* **30(3)**: 1434-1475.

Browning M, Bourguignon F, Chiappori P, Lechene V. 1994. Income and outcomes: a structural model of intrahousehold allocation. *Journal of Political Economy* **102(6)**: 1067-1096.

Browning M, Lechene V. 2001. Caring and sharing: tests between alternative models of intra-household allocation. *Discussion Paper* **01-07**. Department of Economics, Oxford University and Institute of Economics, University of Copenhagen.

Browning M., Chiappori, P, and Lewbel A. 2013 Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power. *Review of Economic Studies*, **80**: 1267-1303.

Burgess R, Zhuang J. 2003. Modernisation and son preference. *LSE STICERD Research Paper* No. DEDPS29.

Case A, Deaton A. 2002. Consumption, health, gender and poverty. Research Program in Development Economics, Princeton University.

Case A, Lin I, McLanahan S. 2000. How hungry is the selfish gene? *Economic Journal* **110 (Oct.)**: 781-804.

Clark A. 2000. Son preference and sex composition of children: evidence from India. *Demography* **3(1)**: 95-108.

Cramer, J.C. 1980. Fertility and Female lyment: Problems of Causal Direction. *American Sociological Review* 45(2): 167-190

Das Gupta M. 1987. Selectivity Discrimination against Female Children in Rural Punjab, India. Population and Development Review. 13(1): 77-99.

Deaton A, Paxson C. 1998. Economy of scale, household size, and the demand for food. *Journal of Political Economy* **106(5)**: 897-930.

Deaton A. 1989. Looking for boys-girls discrimination in household expenditure data. *World Bank Economic Review* **3**: 1-15.

Deaton A., Ruiz-Castillo J., Thomas D. 1989. The influence of household composition on household expenditure patterns. *Journal of Political Economy* **97(1)**: 179-200.

Deaton A. 1987. The allocation of goods within the household: adults, children, and gender. Living Standard Measurement Unit, The World Bank; Research Program in Development Studies, Princeton University.

Deaton A. 1997. The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Johns Hopkins University Press: New York

Deaton A, Muellbauer J. 1986. On measuring child costs: with application to poor countries. *Journal of Political Economy* **94(4)**: 720-744.

Deaton A, Muellbauer J. 1980. An almost ideal demand system. *American Economic Review* **70**: 312-336.

Desta M, Haddis G, Atakelt S. 2006. *Female-headed Households and Livelihood* Intervention in Four Selected Weredas in Tigray, Ethiopia. Report No. **44**, Drylands Coordination Group, Addis Ababa.

Dickens R, Fry V, Pashardes P. 1993. Non-linearities and equivalence scales. *Economic Journal* **103:** 359–368.

Dunbar, G.R., A. Lewbel, and K. Pendakur. 2013. Children Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi, *American Economic Review*, **103 (1)**: 438-471

Fan J, Gijbels I. 1996. *Local Polynomial Modelling and Its Applications*. Chapman & Hall: London.

Fan J. 1992. Design-adaptive nonparametric regression. *Journal of the American Statistical Association* **87(420)**: 998-1004.

Federal Democratic Republic of Ethiopia. 2001. *Report on the 1999/2000 Household Income, Consumption, and Expenditure Survey*. General Statistical Authority: Addis Ababa.

Gibson J, Rozelle S. 2004. Is it better to be a boy? A disaggregated outlay equivalent analysis of gender bias in Papua New Guinea. *Journal of Development Studies* **40(4)**: 115-136.

Gilligan D, Hoddinott J, Kumar N, Taffesse AS. 2009. Can social protection work in Africa? Evidence on the impact of Ethiopia's productive safety net programme on food security, assets and incentives. IFPRI: Washington, DC.

Gong X, von Soest A, Zhang P. 2005. The effects of gender of children on expenditure patterns in rural China: a semiparametric analysis. *Journal of Applied Econometrics* **20**: 509-527.

Gorman WM. 1980. Some Engel curves. In *Essays in Theory and Measurement of Consumer Behaviour*, Deaton A (ed). Cambridge University Press: Cambridge.

Gorman WM. 1976. Tricks with utility functions. In *Essays in Economic Analysis*, Artis M, Nobay AR (eds), pp. 211-243. Cambridge University Press: Cambridge.

Gozalo PL. 1997. Nonparametric bootstrap analysis with applications to demographic effects in demand functions. *Journal of Economterics* **81**: 357-393.

Gronau R. 1991. Intrafamily allocation of goods: how to separate the adult from the child. *Journal of Labor Economics* **9(3)**: 207-235.

Gronau R. 1988. Consumption technology and the intrafamily distribution of resourses: adult equivalence scales reexamined. *Journal of Political Economy* **81(Suppl.)**: S168-S199.

Hardle W, Mammen E. 1993. Comparing nonparametric versus parametric regression fits. *Annals of Statistics* **21** (**4)**, 1926-1947.

Holly A, Sragan J. 1982. Testing for exogeneity in a limited information framework. *American Economic Review* **70**: 263-272.

Jensen, R. 2002. Equal Treatment, Unequal Outcomes? Generating Sex Inequality Through Fertility Behavior. Kennedy School of Government Discussion Paper, Harvard University.

Kebede B. 2008. Intrahousehold allocation in rural Ethiopia: a demand system approach. *Review of Income and Wealth* **54(1)**: 1-26.

Koohi-Kamli F. 2008. Intrahousehold inequality and child gender bias in Ethiopia. *World Bank Policy Research Working Paper* No. 4755. The World Bank.

Koohi-Kamli F. 2016. Gender Bias in Orphanhood: Ethiopian Evidence from Adults Consumption Patterns, *fkoohi.com.*

Kynch J, Sen AK. 1983. Indian women: well-being and survival. *Cambridge Journal of Economics* **7(3/4)**: 363-380.

Lanjouw P, Ravallion M. 1993. Are larger households really poorer? Policy Research Department, World Bank.

Leser CEV. 1963. Forms of Engel functions. *Econometrica* **31(4)**: 694-703.

Leung, SF. 1991. A stochastic dynamic analysis of parental sex preferences and fertility. *Quarterly Journal of Economics* **106(4)**: 1063-1088.

Lyssiotou P, Pashardes P, Stengos T. 2008. Demographic versus expenditure flexibility in Engel curves. *Empirical Economics* **34**: 257-271.

Muellbauer J. 1987. Professor Sen on the standard of living. In *The Standard of Living*, Sen AK, Hawthorn G (eds), pp. 39-58. Cambridge University Press: Cambridge.

Muhuri P, Preston S. 1991. Effects of family composition on mortality differentials by sex among children in Matlab, Bangladesh. *Population and Development Review* **13 (3)**: 415-434.

Nelson JA.1992. Methods of estimating household equivalence scales: an empirical investigation. *Review of Income and Wealth* **38(3):** 295-310.

Nelson JA. 1988. Household economies of scale in consumption: theory and evidence. *Econometrica* **56(6):** 1301-1314.

Pendakar K. 1999. Semiparametric estimation and tests of base-independent equivalence scales. *Journal of Economterics* **88**: 1-40.

Robinson PM. 1988. Root-N-consistent semiparametric estimation. *Econometrica* **56**: 931-954.

Rothbarth E. 1943. Note on a method of determining equivalent income for families of different composition. Appendix 4 in *War-time Pattern of Saving and Spending*, Madge C. Cambridge University Press: Cambridge.

Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: London.

Subramanian S., Deaton A. 1991. Gender effects in Indian consumption patterns. *Sarvekshana* **14**: 1-12.

Working H. 1943. Statistical laws of family expenditure. *Journal of the American Statistical Association* **38**: 43-56.

Wu CFJ. 1986. Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics* **14**: 1261-1350.

Yatchew A, No A. 2001. Household gasoline demand in Canada. *Econometrica* **69**: 1697-1709.

Young Lives 2014. Ethiopia-Nutrition-FactSheet, Nutrition and Health: Round 4 Preliminary Findings,

Appendix A-Adult goods Budget Share estimates with main controls

Tables 1a-3a present the results given the main text with a more extensive set of controls used; additional controls such as regional variables, etc. were also employed but not reported here to save space (the full results are available from the author.) The main points to note are that child age effects, and the effect of rural-urban dummy; employment safety and school feeding program dummies are often very significant in all three tables. Moreover, there are significant curvature effects in table 1a with expenditure endogeneity control but a similar effect is only observable for personal item when we control for household size endogeneity in table 2a and for coffee in table 3a.

Table 1a Semiparametric and parametric estimation of child gender bias effects on the budget shares of four principal adult goods (*t*-ratios in brackets) with **control for total expenditure endogeneity**

Dep. Var	Personal Services **change for new test**		Adult clothing		Tobacco **change for new test**		Coffee **change for new test**	
Variable	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet
ln (*Y/n)*	_	0.0137	_ **OK**	0.0133	_	-0.0105	_	0.0357
		(3.60)		(1.99)		(-1.45)		(2.17)
ln $(Y/n)^2$	_	-0.0010	_	-0.0004	_	0.0009	_	-0.0019
		(-6.08)		(-1.68)		(2.28)		(-2.41)
(girls/n)	**0.0025**	**0.0024**	**0.0020**	**0.0021**	**0.0018**	**0.0018**	**0.0041**	**0.0040**
	(5.25)	**(5.18)**	**(2.03)**	**(2.05)**	**(2.37)**	**(2.80)**	**(2.38)**	**(2.36)**
ln *n*	0.0004	0.0004	0.0038	0.0038	0.0013	0.0013	0.0064	0.0064
	(0.39)	(0.39)	(1.80)	(1.83)	(0.58)	(1.01)	(1.87)	(1.87)
ch0-2/n	-0.0003	-0.0003	-0.0018	-0.0017	0.0021	0.0022	0.0004	0.0003
	(-0.31)	(-0.39)	(-1.81)	(-1.74)	(1.72)	(2.08)	(0.12)	(0.10)
ch3-6/n	-0.0027	-0.0028	-0.0015	-0.0014	0.0020	0.0020	0.0010	0.0011
	(-2.70)	(-2.73)	(-1.13)	(-1.10)	(1.75)	(1.88)	(0.56)	(0.59)
ch7-9/n	-0.0010	-0.0010	-0.0023	-0.0023	0.0003	0.0004	-0.0004	-0.0007
	(-1.15)	(-1.16)	(-1.53)	(-1.53)	(0.25)	(0.40)	(-0.16)	(-0.31)
ch10-12/n	0.0011	0.0011	-0.0017	-0.0016	0.0011	0.0010	0.0040	0.0037
	(0.71)	(0.69)	(-0.94)	(-0.94)	(0.55)	(0.62)	(1.07)	(0.98)
rural dummy	-0.0009	-0.0010	-0.0028	-0.0028	-0.0011	-0.0010	-0.0090	-0.0094
	(-1.31)	(-1.39)	(-2.16)	(-2.27)	(-0.73)	(-0.79)	(-3.78)	(-3.91)
SNP safety	0.0022	0.0022	0.0001	0.0001	-0.0046	-0.0046	0.0019	0.0019
net	(2.78)	(2.83)	(0.11)	(0.07)	(-6.85)	(-3.49)	(0.44)	(0.45)
school	-0.0005	-0.0005	0.0030	0.0030	0.0024	0.0025	-0.0054	-0.0053
feeding	(-0.54)	(-0.59)	(1.19)	(1.19)	(2.33)	(2.38)	(-1.80)	(-1.76)
R^2	-	0.0318	-	0.0165	-	0.1316	-	0.2487
endogeneity	0.0014	0.0014	-0.0066	-0.0069	-0.0030	-0.0031	-0.0085	-0.0086
control ρ_x	(0.50)	(0.49)	(-1.26)	(-1.30)	(-0.48)	(-0.75)	(-0.72)	(-0.74)
mean depVar	0.0039	0.0039	0.0023	0.0023	0.0128	0.0128	0.0128	0.0128

Table 2a Semiparametric and parametric estimation of child gender bias effects on the budget shares of four principal adult goods (*t*-ratios in brackets) with **control number of children endogeneity**

Dep. Var	Personal Services		Adult clothing		Tobacco		Coffee	
Variable	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet
ln (*Y/n)*	-	0.0151	_	0.0247	_	-0.0098	_	-0.0556
		(5.04)		(1.56)		(-0.55)		(-1.08)
ln $(Y/n)^2$	_	-0.0010	_	-0.0014	_	0.0007	_	0.0025
		(-6.09)		(-1.59)		(0.69)		(0.91)
(girls/n)	**0.0025**	**0.0025**	**0.0022**	**0.0022**	**0.0014**	**0.0015**	**0.0278**	**0.0277**
	(4.09)	**(4.04)**	**(2.42)**	**(2.40)**	**(2.05)**	**(2.44)**	**(1.66)**	**(1.64)**
ln *n*	0.0008	0.0006	0.0176	0.0168	0.0051	0.0039	-0.1102	-0.1098
	(0.11)	(0.09)	(1.39)	(1.30)	(0.57)	(0.27)	(-1.47)	(-1.45)
ch0-2/n	-0.0003	-0.0004	0.0058	0.0055	0.0046	0.0040	0.0151	0.0148
	(-0.35)	(-0.44)	(0.98)	(0.91)	(0.87)	(0.46)	(1.72)	(1.68)
ch3-6/n	-0.0028	-0.0028	-0.0028	-0.0027	0.0021	0.0020	0.0484	0.0483
	(-2.23)	(-2.24)	(-1.92)	(-1.83)	(1.29)	(0.89)	(1.57)	(1.55)
ch7-9/n	-0.0012	-0.0011	-0.0062	-0.0060	-0.0003	-0.00006	0.0387	0.0381
	(-0.52)	(-0.50)	(-2.28)	(-2.15)	(-0.16)	(-0.03)	(1.46)	(1.42)
ch10-12/n	0.0009	0.0010	-0.0055	-0.0053	-0.0007	-0.0002	0.0160	0.0159
	(0.18)	(0.20)	(-1.79)	(-1.68)	(-0.17)	(-0.04)	(1.84)	(1.82)
rural dummy	-0.0012	-0.0013	0.0026	0.0024	-0.0027	-0.0028	-0.0337	-0.0341
	(-1.09)	(-1.17)	(1.08)	(0.99)	(-1.85)	(-1.13)	(-3.03)	(-3.04)
SNP safety	0.0023	0.0023	-0.0005	-0.0005	-0.0056	-0.0057	0.0029	0.0028
net	(2.91)	(2.96)	(-0.77)	(-0.77)	(-7.17)	(-3.47)	(0.69)	(0.69)
school	-0.0004	-0.0004	0.0029	0.0028	0.0023	0.0024	-0.0141	-0.0142
feeding	(-0.43)	(-0.48)	(1.30)	(1.27)	(2.27)	(2.01)	(-3.70)	(-3.75)
R^2	-	0.0318	-	0.0169	-	0.1124	-	0.2261
endogeneity	0.00003	0.0002	-0.0165	-0.0158	-0.0050	-0.0037	0.1140	0.1137
control ρ_c	(0.00)	(0.03)	(-1.28)	(-1.20)	(-0.55)	(-0.25)	(1.51)	(1.49)
mean depVar	0.0039	0.0039	0.0023	0.0023	0.0128	0.0128	0.0128	0.0128

Table 3a Semiparametric and parametric estimation of child gender bias effects on budget shares of four principal adult goods (*t*-ratios in brackets) **control for children endogeneity by logit residual**

Dep. Var	Personal Services		Adult clothing		Tobacco		Coffee	
Variable	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet	Semipar	Paramet
ln (*Y/n)*	-	0.0141	_	0.0135	_	-0.0107	_	0.0368
		(3.66)		(1.98)		(-1.47)		(2.20)
ln $(Y/n)^2$	_	-0.0010	_	-0.0004	_	0.0009	_	-0.0020
		(-6.13)		(-1.70)		(2.29)		(-2.74)
(girls/n)	**0.0025**	**0.0025**	**0.0021**	**0.0021**	**0.0018**	**0.0018**	**0.0037**	**0.0038**
	(5.27)	**(5.20)**	**(2.06)**	**(2.08)**	**(2.36)**	**(2.83)**	**(2.18)**	**(2.18)**
ln *n*	0.0005	0.0005	0.0038	0.0040	0.0013	0.0013	0.0085	0.0083
	(0.57)	(0.55)	(1.84)	(1.86)	(0.58)	(1.05)	(2.36)	(2.34)
ch0-2/n	-0.0002	-0.0003	-0.0017	-0.0016	0.0021	0.0022	0.0009	0.0008
	(-0.23)	(-0.32)	(-1.71)	(-1.65)	(1.68)	(2.07)	(0.28)	(0.25)
ch3-6/n	-0.0028	-0.0028	-0.0016	-0.0015	0.0019	0.0019	0.0039	0.0005
	(-2.69)	(-2.72)	(-1.17)	(-1.13)	(1.96)	(1.81)	(0.21)	(0.26)
ch7-9/n	-0.0010	-0.0011	-0.0024	-0.0024	0.0003	0.0004	-0.0003	-0.0007
	(-1.18)	(-1.18)	(-1.56)	(-1.55)	(0.19)	(0.33)	(-0.14)	(-0.30)
ch10-12/n	0.0011	0.0011	-0.0017	-0.0017	0.0011	0.0010	0.0038	0.0036
	(0.69)	(0.69)	(-0.97)	(-0.96)	(0.56)	(0.63)	(1.01)	(0.93)
rural dummy	-0.0009	-0.0009	-0.0027	-0.0028	-0.0011	-0.0010	-0.0089	-0.0093
	(-1.20)	(-1.30)	(-2.12)	(-2.22)	(-.73)	(-0.78)	(-3.79)	(-3.90)
SNP safety	0.0022	0.0022	0.0001	0.0001	-0.0046	-0.0047	0.0019	0.0019
net	(2.75)	(2.81)	(0.11)	(0.07)	(-6.86)	(-3.51)	(0.44)	(0.44)
school	-0.0004	-0.0005	0.0030	0.0030	0.0024	0.0025	-0.0053	-0.0052
feeding	(-0.51)	(-0.56)	(1.18)	(1.18)	(2. 32)	(2.38)	(-1.75)	(-1.72)
R^2	-	0.2491	-	0.0165	-	0.1317	-	0.2491
endogeneity	0.0014	0.0014	-0.0067	-0.0069	-0.0029	-0.0031	-0.0079	-0.0081
control ρ_x	(0.49)	(0.48)	(-1.26)	(-1.30)	(-0.46)	(-0.73)	(-0.67)	(-0.69)
endogeneity	0.0009	0.0007	-0.0096	-0.0011	0.0017	0.0019	-0.0039	-0.0034
control ρ_k	(0.70)	(0.60)	(-0.53)	(-0.62)	(0.44)	(0.65)	(-0.81)	(-0.75)
μ_k**ln (Y/n)*	-0.0002	-0.0002	0.0001	0.0001	-0.0002	-0.0003	0.0002	0.0001
	(-0.89)	(0.80)	(0.37)	(0.46)	(-0.43)	(-0.66)	(0.31)	(0.25)
mean depVar	0.0039	0.0039	0.0023	0.0023	0.0128	0.0128	0.0128	0.0128

Notes to Tables 1a-3a: Standard errors are adjusted for clustering. *Y/n* is per capita total expenditure, dep.var. stands for the dependent variables: the budget shares of adult goods and services, adult clothing, tobacco, and coffee. Demographics are log household size *n*, proportions of girls and of children aged 0-2, 3-6, 7-9, and 10-12, leaving those above 12 as the age reference group (a child defined as < 15 years of age). See notes to tables 4 and 6 on semiparametric estimation method and functional form specification, and Appenix B, tables'2a-3c for the first stage estimates of the instruments employed and logit estimates.

Appendix B-First stage estimates by IV and logit residual for expenditure & child number.

Table 1a First-stage Instruments, log. **total expenditure** (abs. t-ratios in brackets)				
	Personal	Cloths	Tobacco	Coffee
Land	0.1114 (4.60)	-0.0048 (1.74)	0.1024 (4.10)	0.1000 (4.10)
Stove	0.0238 (1.58)	-0.0008 (1.03)	0.0261 (1.67)	0.0251 (1.69)
Radio	-0.0214 (1.33)	-0.0006 (1.03)	-0.0241 (1.44)	-0.0239 (1.50)

Notes: Instrumental dummies are for land defined as a 1 if owned, 0 otherwise, stoves (gas/electric) defined as 1 is owned, 0 otherwise; and radio defined as 1 if owned, 0 otherwise.

Table 1b First-stage Instruments for log. **Number of children**, average per warada (t-ratios in brackets)				
	Personal	Cloths	Tobacco	Coffee
Dschool	-0.0068 (2.16)	-0.0048 (1.74)	-0.0070 (2.47)	-0.0008 (0.35)
Tschool	-0.0001 (0.13)	-0.0008 (1.03)	-0.0001 (0.16)	-0.0005 (0.90)
Thealth	-0.0005 (0.78)	-0.0006 (1.03)	-0.0008 (1.32)	-0.0001(0.29)
Dwater	0.0038 (3.42)	0.0025 (2.42)	0.0025 (2.26)	0.0012 (1.34)
Twater	0.0082 (1.16)	0.0095 (1.63)	0.0089 (1.35)	-0.0059 (0.97)
Twood	-0.0022 (2.96)	-0.0013 (1.59)	-0.0021 (2.51)	-0.0012 (2.05)

Notes: Average time per wereda in minutes taken the nearest primary school, health center, sources of drinking water and heat and cooking fire-wood collections; average distance per wereda in km to the nearest primary school and source of drinking water.

Table 2c- Logit estimates of probability of more than one child conditional on presence of oldest boy, or on at least a boy, in the household, *z*-values in brackets.

oldest boy	0.4831 (4.52)	-
at least a boy	-	1.6821 (14.79)
ch0-2/n	7.6594 (16.82)	7.5148 (16.05)
ch3-6/n	8.7091 (19.01)	8.4645 (18.02)
ch7-9/n	7.0285 (15.35)	6.9477 (14.78)
ch10-12/n	0.2233 (0.57)	0.2501 (0.62)
log mother age	34.8545 (7.53)	34.5426 (7.18)
$(\log \text{mother age})^2$	-5.2434 (-7.80)	-5.1807 (-7.47)
mother educ: basic literacy	0.9838 (4.34)	0.9499 (3.98)
mother educ: highest grade	0.0158 (2.77)	0.0160 (2.76)
$\ln (Y/n)$	-3.0809 (-2.75)	-3.0815 (-2.66)
$\ln (Y/n)^2$	0.1707 (2.32)	0.1701 (2.24)
rural dummy	0.0100 (0.08)	-0.0312 (-0.23)
SNP safe net	0.2204 (1.56)	-0.3633 (-1.79)
school feed	-0.2954 (-1.50)	0.2489 (1.68)

*Explanatory variables include work and marital status of the head and spouse; age, age^2 and education of the head, and nine regional dummies (reference Addis Ababa)