American Economic Journal: Applied Economics 2012, 4(2): 1–35
http://dx.doi.org/10.1257/app.4.2.1

What Linear Estimators Miss: The Effects of Family Income on Child Outcomes†

By Katrine V. Løken, Magne Mogstad, and Matthew Wiswall*

We assess the implications of nonlinearity for IV and FE estimation when the estimated model is inappropriately assumed to be linear. Our application is the causal link between family income and child outcomes. Our nonlinear IV and FE estimates show an increasing, concave relationship between family income and children's outcomes. We find that the linear estimators miss the significant effects of family income because they assign little weight to the large marginal effects in the lower part of the income distribution. We also show that the linear IV and FE estimates differ primarily because of different weighting of marginal effects. (*JEL* C26, D14, J12, J13)

Many recent empirical papers seek to estimate causal relationships using instrumental variables (IV) or fixed effects (FE) methods when concerns about endogeneity bias arise. When the potentially endogenous regressor takes on several values, the standard approach is to specify a *linear* model of the following type:

$$y = \mu + \beta C + \mathbf{X}'\delta + \epsilon,$$

where y is the dependent variable; C is the potentially endogenous regressor: and $\mathbf{X}$ is a vector of exogenous covariates. While many empirical and econometric studies explore the consequences for IV and FE estimation of parameter heterogeneity when C is binary,[1] very few studies focus on the implications of nonlinearity when C is multivalued and the estimated model is assumed to be linear. Angrist and Imbens (1995); Angrist, Graddy, and Imbens (2000); Lochner and Moretti (2001, 2011); and Mogstad and Wiswall (2011a, b) are notable exceptions that explore the implications of inappropriately assuming linearity in IV estimation.

* Løken: University of Bergen, Department of Economics, Postboks 7802, 5020 Bergen, Norway (e-mail: katrine.loken@econ.uib.no); Mogstad: University College London, Department of Economics, Gower Street, London WC1E 6BT, United Kingdom (e-mail: m.mogstad@ucl.ac.uk); Wiswall: New York University, Department of Economics, 19 W. 4th St., New York, NY 10012 (e-mail: matt.wiswall@gmail.com). The Norwegian Research Council has provided financial support for this project. Thanks to Arild Aakvik, Gordon Dahl, Monique De Haan, Kjell Lommerud, and two anonymous referees for helpful comments and suggestions.

† To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://dx.doi.org/10.1257/app.4.2.1.

[1] Influential work by Imbens and Angrist (1994) and Heckman and Vytlacil (2005) has clarified the interpretation of IV estimates as local average treatment effects when the regression parameter of interest is heterogenous.

Yet in many empirical applications, there is no particular reason to expect that the true relationship is linear.[2]

The aim of our paper is to examine the implications of nonlinearity for IV and FE estimation of the causal link between childhood family income and child outcomes.[3] Although the linear model used in previous studies is convenient and may be preferable on grounds of efficiency, it runs counter to economic theory predicting an increasing concave relationship. In particular, the seminal model of Becker and Tomes (1979, 1986) suggests that low-income families will underinvest in their children's human capital because they are more likely to be credit constrained. The marginal return on human capital will therefore exceed that on assets in credit constrained families. This implies that poor parents will invest more of an increase in family income in children's human capital compared to rich parents who will give more as bequests. In order to test this theory and assess how poverty affects a child's human capital development, it is necessary to move beyond linear estimates.[4]

We begin by clarifying the interpretation of OLS and IV estimators in the case where the true model is nonlinear, but the researcher estimates a linear model. The linear OLS and IV estimators will then identify different weighted averages of the underlying marginal causal effects, even in the absence of endogeneity or heterogeneity in the marginal effects (see Angrist and Imbens 1995, Angrist and Krueger 1999; and Mogstad and Wiswall 2011a). To our knowledge, however, such an identification result does not exist for the linear FE estimator. We therefore show how the linear FE estimator identifies a weighted average of the underlying marginal causal effects. Just as for the OLS and IV weights, the FE weights have an intuitive interpretation, are functions of observable quantities, and can be estimated under very general assumptions. But because the FE weights differ from the OLS/IV weights, inappropriately assuming linearity will generally yield different FE and OLS/IV estimates, even in the absence of endogeneity or heterogeneity in the marginal effects. We conclude the methodological part of our paper by showing how a Blinder-Oaxaca type of decomposition method can be used to measure the contribution of different weights to the differences between linear OLS, FE, and IV estimates.

These methodological insights motivate and guide our empirical analysis of the effects of childhood family income on child development. Despite a large body of evidence showing a positive association between family income and child development, there is much controversy about whether these correlations can be given causal interpretations. Unobservable determinants of children's outcomes that are correlated with family income, like parental abilities, are of major concern when assessing the causal impact of family income on child development. While most

[2] A prominent example is the economic returns to schooling, where a considerable body of evidence suggests a nonlinear relationship between log earnings and schooling (Heckman 2008). Yet, the log-linear model dominates the empirical literature that uses IV or (twin) FE methods.

[3] For reviews of the empirical literature on childhood family income and child outcomes, see Mayer (1997), Solon (1999), and Almond and Currie (2010).

[4] As emphasized by Heckman (2008), there are two distinct types of credit constraints operating on the family and its children. The first constraint is the inability of parents to borrow against their children's future income to finance investment in them. The second constraint is the inability of parents to borrow against their own income to finance investment in their children. Both types of constraints can produce a nonlinear relationship between family income and child outcomes.

previous studies have used family-specific FE estimators to eliminate biases from permanent family characteristics, a recent strand of the literature exploits quasi-natural experiments to instrument for family income. Although these studies represent a significant step forward, the evidence is far from conclusive. While some studies report small and sometimes insignificant effects of family income on child outcomes, others suggest substantial positive effects.[5] Our empirical analysis shows how relaxing the linearity restriction in the FE and IV estimation changes the qualitative conclusions about the effects of family income on child outcomes, and can be important to reach a consensus about the causal link between family income and child outcomes.

We use administrative registers for the entire population of Norway, with information on children's educational attainment and IQ as adults as well as their family income during childhood. To instrument for family income, we exploit regional and time variation in the economic boom that followed the initial discovery of oil in Norway as the instrument for family income. In doing so, we are able to control for unobserved permanent differences between children born in different years as well as between children born in different areas.[6]

Our linear IV and FE results show estimates of family income on children's adult IQ and educational attainment that are insignificant and much smaller than the sizable and significant OLS estimates. In the spirit of previous studies, our results could be interpreted as suggesting little, if any, causal effect of family income on children's outcomes, once endogeneity bias is addressed by IV or FE methods. However, when relaxing the linearity restriction in family income, the IV and FE estimates line up with the theoretical prediction of an increasing, concave relationship. The evidence of large marginal effects in the lower part of the distribution indicates that income support programs targeted at poor families might be quite effective in promoting child development.

To understand why the linear FE and IV estimates miss the significant impact of family income on child outcomes, we estimate the set of weights for each estimator. We find that the linear IV estimate assigns little weight to the large marginal effects in the lower part of the income distribution, reflecting that the oil boom did not do much for family income of poor families. We also find that poor families in Norway experienced little within-family income variation, implying that small marginal effects in the middle and upper part of the income distribution are weighted heavily in the linear FE estimate. In comparison, the sizable linear OLS estimate assigns much more weight to marginal effects in the lower part of the income distribution.

[5]While the IV estimates reported in Oreopoulos et al. (2008), Dahl and Lochner (2011), and Milligan and Stabile (2007) suggest some positive effects of family income on children's (short-run) outcomes, Shea (2000), and Løken (2010) find little, if any, impact of family income. Using FE estimation, both Duncan et al. (1998), and Levy and Duncan (2000) find that family income is important for children's educational attainment, whereas Blau (1999), and Dooley and Stewart (2004) find a small effect of family income on children's outcomes.

[6]In the Norwegian context, cash benefits to families with children may reduce the importance of parents' constraints on borrowing against their own income to finance investment in their children. However, the welfare system in the period we consider was much less generous than today. Moreover, cash benefits do not relax the child's constraint on borrowing against his or her future income. Credit constraints may, therefore, produce a nonlinear relationship between family income and child outcomes, even in Norway.

We next examine how much of the difference across the linear estimates the weights explain. With homogeneity in the marginal effects (and no measurement error), a consistent FE estimate may differ from a consistent IV estimate only because of different weighting (and sampling error). We expect, however, the OLS estimates to be (upwardly) biased, in which case the linear OLS estimates may differ also due to different marginal effect estimates. Our decomposition method reveals that the IV estimates exceed the FE estimates primarily because the latter assign less weight to the sizable marginal effects in the middle part of the distribution. In comparison, differences in the marginal effect estimates are clearly an important factor behind the relatively high OLS estimates.

The remainder of the paper is organized as follows. Section I shows what linear OLS, IV, and FE estimators identify, and describes the possible implications of nonlinearity. Section II describes our data and discusses the natural experiment used as an instrument for family income. Section III presents the empirical results, and Section IV summarizes and concludes with a discussion of the general lessons that can be drawn from our study.

I. What Linear Estimators Identify

In this section, we discuss the implications of nonlinearity for the interpretation and comparison of linear OLS, IV, and FE estimators. For simplicity, this section ignores control variables, but we will include them in the empirical analysis.

A. *Potential Outcomes, Linearity, and Marginal Effects*

Let $f_i(c)$ denote the potential (or latent) outcome that child i would receive with level c of childhood family income. In the context of a theoretical model of the relationship between family income and child outcome, the functional form of $f_i(c)$ may be determined by aspects of individual behavior and/or market forces, like in Becker and Tomes (1979, 1986). With or without an explicit theoretical model for $f_i(c)$, however, we can think of this function as describing the outcome level that child i would achieve if he or she was assigned childhood family income c (e.g., in an experiment).

The observed level of family income for child i is denoted by C_i. The standard regression framework used in the literature to link the potential outcome to the observed outcome, y_i, has the following form:

$$y_i = f_i(C_i) = \mu + \beta C_i + \epsilon_i, \tag{1}$$

where ϵ_i is a mean-zero error term, which captures unobserved factors determining child outcomes. This model forms the basis for previous studies using IV and FE methods to examine the impact of family income on child outcomes.

Our point of departure is to relax the linearity assumption and allow the marginal effects on children's outcomes of an increase in family income to vary across the family income distribution, as theory suggests. Let family income take on values in the finite set $C_i \in \{0, 1, \ldots, \bar{c}\}$. Using dummy variables constructed as $d_{ci} = 1\{C_i \geq c\}$, we can specify an unrestricted model in family income:

$$y_i = \mu + \sum_{c=1}^{\bar{c}} \gamma_c d_{ci} + \upsilon_i, \tag{2}$$

where υ_i is a mean-zero error term, and the γ_c coefficient represents the marginal effect of a one unit (e.g., USD 1) increase from family income level $c - 1$ to c. The unrestricted model (2) nests the linear model (1), which restricts the marginal effects to be independent of family income level: $\gamma_c = \beta$ for all $c > 0$.

B. *OLS Decomposition*

The OLS estimand for β in (1) is $\beta(OLS) = \text{Cov}(y_i, C_i)/\text{Var}(C_i)$. As shown in Mogstad and Wiswall (2011a), we can write the linear OLS estimand as

$$\beta(OLS) = \sum_{c=1}^{\bar{c}} \gamma_c(OLS)\, w_c(OLS), \tag{3}$$

where

$$\gamma_c(OLS) = E[y_i | C_i = c] - E[y_i | C_i = c - 1] = \gamma_c + \Delta_c,$$

and $\Delta_c = E[\upsilon_i | C_i = c] - E[\upsilon_i | C_i = c - 1]$ represents the selection bias in the OLS estimates of the marginal effects, $\gamma_c(OLS)$. The associated OLS weight on $\gamma_c(OLS)$ is defined as

$$w_c(OLS) = \frac{\text{Cov}(d_{ci}, C_i)}{\text{Var}(C_i)}.$$

These OLS weights are simply the regression coefficient of $d_{ci} = 1\{C_i \geq c\}$ on C_i. The OLS weights sum to one, are nonnegative, and can be directly estimated using the sample analog of the above expressions.[7]

There are two key issues with OLS estimation of (1). As is well known, OLS estimates will be biased if observed family income is correlated with the unobserved factors determining child outcomes. The second issue, which has received far less attention in empirical research, is that the linear OLS estimand has a particular weighting over the marginal effects, given by (3). Specifically, weight is given to each γ_c in proportion to the conditional mean of C_i, above and below C. More weight is also given to marginal effects close to the sample median of C_i, since this is where $\Pr(C_i \geq c)(1 - \Pr(C_i \geq C))$ is maximized. If there are nonlinearities in the OLS estimates of the marginal effects ($\gamma_c(OLS) \neq \gamma_{c'}(OLS)$ for $c \neq c'$), then it follows from (3) that the linear OLS estimate depends on how it weights the marginal effects, and thereby the sample distribution of C_i.

[7]Note that the decomposition of the sample analog of $\beta(OLS)$ has exactly the same form: $\hat{\beta}(OLS) = \sum_{c=1}^{\bar{c}} \hat{w}_c(OLS)\hat{\gamma}_c(OLS)$, where $\hat{\beta}(OLS)$, $\hat{w}_c(OLS)$, and $\hat{\gamma}_c(OLS)$ are the sample analogs of the estimands.

C. *IV Decomposition*

Angrist and Imbens (1995) provide an analogous decomposition of the linear IV estimand in the case of a binary instrument and under an assumption of monotonicity. As our empirical analysis uses a multi-valued instrument, we use a generalized version of the decomposition of the linear IV estimand, provided in Mogstad and Wiswall (2011a).

Consider the case of a scalar (binary or multi-valued) instrument Z_i. Suppose that the standard IV assumptions hold:

ASSUMPTION 1 (*Instrument Uncorrelated with Residual*): $\mathrm{Cov}(\upsilon_i, Z_i) = 0$.

ASSUMPTION 2 (*Instrument Correlated with Family Income*): $\mathrm{Cov}(C_i, Z_i) \neq 0$.

These assumptions imply that the instrument is uncorrelated with the unobserved factors determining child outcomes, and that the instrument has some effect on family income.

Under Assumptions 1 and 2, the linear IV estimand for β in (1) is $\beta(z) = \mathrm{Cov}(y_i, Z_i)/\mathrm{Cov}(C_i, Z_i)$. Mogstad and Wiswall (2011a) show that $\beta(z)$ can be decomposed as

$$\beta(Z) = \sum_{c=1}^{\bar{c}} w_c(Z)\gamma_c, \tag{4}$$

where

$$w_c(Z) = \frac{\mathrm{Cov}(d_{ci}, Z_i)}{\mathrm{Cov}(C_i, Z_i)}.$$

These weights sum to one, and can be computed using the sample analog of the above expression.[8] The weights will be nonnegative if the monotonicity assumption holds, that is, the instrument affects everyone in the same way, if at all (Angrist and Imbens 1995).

From (4), we learn that $\beta(Z)$ is a weighted average of the marginal effects across the family income distribution. The weight $w_c(Z)$ attached to γ_c depends on the proportion of children who, because of the instrument, experience a change in family income from less than c to c or more. Hence, $\beta(Z)$ assigns more weight to the marginal effects for the levels of family income that are most affected by the particular instrument chosen.

One important feature of the linear IV estimand is that the weights $w_c(Z)$ are functions of the chosen instrument Z, implying that other instruments can lead to different weights and different $\beta(Z)$. If there are nonlinearities in the marginal effects ($\gamma_c \neq \gamma_{c'}$ for $c \neq c'$), linear IV estimators based on different instruments will generally produce disparate estimates of the effect of family income. Hence, previous

[8]Note that the decomposition of the sample analog of $\beta(Z)$ has exactly the same form: $\hat{\beta}(Z) = \sum_{c=1}^{\bar{c}} \hat{w}_c(Z)\hat{\gamma}_c(Z)$, where $\hat{\beta}(Z)$, $\hat{w}_c(Z)$, and $\hat{\gamma}_c(Z)$ are the sample analogs of the estimands.

studies may have reached conflicting conclusions about the effect of family income on child outcome because their linear IV estimates capture marginal effects at different parts of the family income distribution. For example, the fact that Dahl and Lochner (2011) report substantial positive effects of family income, whereas Løken (2010) finds little if any effect, may simply be because of nonlinearities. While the Earned Income Tax Credit welfare reform instrument used in the former study primarily changed the family income of relatively poor families, the oil boom instrument used in the latter study had the largest impact on the middle and upper part of the family income distribution.

D. *FE Decomposition*

We now move to developing a novel decomposition of the linear FE estimator. This decomposition allows us to understand how the FE estimator weights the various marginal effects and directly compares the linear FE weighting to that of the linear OLS and IV estimators.

As the FE estimator requires information on outcomes and family income for pairs of siblings, we need to extend the cross-sectional setup used in the discussion of the OLS and IV estimators to a panel data setting. Let C_{jb} denote observed childhood family income of sibling b from family j, which as above is assumed to take on values in the finite set $C_{jb} \in \{0, 1, \ldots, \bar{c}\}$. Let $f_{jb}(c)$ represent the potential outcome that child b from family j would receive with level c in childhood family income. For simplicity, we focus our attention on the two sibling FE estimator, so that $b \in \{1, 2\}$.

The motivation for FE estimation is the concern that OLS estimates could be biased because of some fixed unobserved family characteristic correlated with family income and child outcomes, such as inheritable parental characteristics. Suppose that siblings share a common family-specific fixed effect, μ_j, which is potentially heterogeneous in the population and possibly correlated with the level of family income. The potential outcome for sibling $b \in \{1, 2\}$ from family j can then be linked to the observed outcome, $f_{jb}(C_{jb})$, in the following way:

$$y_{jb} = f_{jb}(C_{jb}) = \mu_{jb} + \sum_{c=1}^{\bar{c}} \gamma_c d_{cjb} + v_{jb}, \tag{5}$$

where $d_{cjb} = 1\{C_{jb} \geq c\}$ and $\mu_{jb} = \mu_j + \alpha_b$, where μ_j is the family-specific fixed effect and α_b is a sibling-specific $b \in \{1, 2\}$ intercept. Without loss of generality, we normalize $\alpha_1 = 0$, implying that $\mu_{1j} = \mu_j$ and $\mu_{2j} = \mu_j + \alpha_2$. Restricting the siblings to have the same intercept is of course a special case of (5), where $\alpha_2 = 0$.[9]

The linear FE model restricts the marginal effects of additional income to be constant across the family income distribution, $\gamma_c = \beta$ for all c, and is given by

$$y_{jb} = \mu_{jb} + \beta C_{jb} + \epsilon_{jb}. \tag{6}$$

[9] In a two period panel model, the α_2 term allows for a time-specific fixed effect. For example, if siblings were ordered by sequence of birth in the first difference transformation, α_2 would be a birth order fixed effect.

As is well known, the family-specific fixed effects can be eliminated by taking a difference transformation of (6) between all pairs of siblings, which yields the first-differences model

$$\Delta y_j = \alpha_2 + \beta \Delta C_j + \Delta \epsilon_j, \tag{7}$$

where $\Delta y_j = y_{j2} - y_{j1}$, $\Delta C_j = C_{j2} - C_{j1}$, and $\Delta \epsilon_j = \epsilon_{j2} - \epsilon_{j1}$ are the sibling difference in outcome, family income during childhood, and the residual. For notational purposes, and without loss of generality, we sort siblings by their family income before taking the difference transformation, so that $\Delta C_j \geq 0$ for all j. This implies that $\Delta C_j \in \{0, 1, \ldots, \bar{c}\}$, so that both family income levels and family income changes take on the same possible values.

The linear FE estimand for β in (6) is $\beta(FE) = \mathrm{Cov}(\Delta y_j, \Delta C_j)/\mathrm{Var}(\Delta C_j)$, and can be obtained by performing OLS on (7). Let $\Delta v_j = v_{j2} - v_{j1}$. In line with the previous literature, we consider the following FE assumptions:

ASSUMPTION 3 (*Mean-Independence of Family Income Variation*): $E[\Delta v_j \mid \Delta C_j] = 0$.

ASSUMPTION 4 (*Existence of Family Income Variation*): $\mathrm{Var}(\Delta C_j) \neq 0$.

Assumption 3 implies that the differences between siblings in unobservables are uncorrelated with the differences in family income during their childhood. Assumption 4 is satisfied if there is some variation between siblings in their childhood family income. Under these assumptions, the sample analog of $\beta(FE)$ provides a consistent estimate of β.

Given that the FE estimator is simply an OLS estimator on differences in income, we could of course apply the OLS decomposition on the first-differences model. However, such a decomposition would not be that useful, because it could not reveal how the linear FE estimator weights the marginal effects in family income *levels*. The reason is that a given difference in family income can occur at different income levels, and, therefore, each marginal effect in family income differences is a weighted average of the various marginal effects in levels of family income $\gamma_1, \ldots, \gamma_{\bar{c}}$.

To derive a FE decomposition in terms of marginal effects in levels of family income, we exploit that the linear FE estimator can be expressed as a special case of the linear IV estimator using the instrument q_{jb}, defined for the pooled sample of siblings as:

$$q_{jb} = 1\{b = 2\}\tilde{\Delta}C_j - 1\{b = 1\}\tilde{\Delta}C_j,$$

where $\tilde{\Delta}C_j = \Delta C_j - (1/J)\sum_{j=1}^{J} \Delta C_j$, and J denotes the number of families. In the Appendix, we show that the linear IV estimator using the instrument q_{jb} is numerically equivalent to the linear FE estimator. The q_{jb} instrument is the deviation of the change in family income for each sibling from the mean change in family income $(1/J)\sum_{j=1}^{J}\Delta C_j$. The linear FE weighting on the marginal effects γ_c can therefore be expressed as

$$\beta(FE) = \sum_{c=1}^{\bar{c}} w_c(FE)\gamma_c, \tag{8}$$

where

$$w_c(FE) = \frac{\mathrm{Cov}(d_{cjb}, q_{jb})}{\mathrm{Cov}(C_{jb}, q_{jb})}.$$

These weights sum to one, and can be computed using the sample analog of the above expression.[10]

The FE decomposition is quite general, and can be applied to any two-period panel data setting which satisfies Assumptions 3 and 4. Marginal effects for family income levels that experience most of the within-family income variation receive the most weight in the linear FE estimand. If there are nonlinearities ($\gamma_c \neq \gamma_c$ for $c \neq c'$), then it follows that $\beta(FE)$ depends on how the marginal effects are weighted, and thereby the sample distribution of within-family income variation.

One implication of the dependence of $\beta(FE)$ on the sample distribution of within-family income variation, is that the linear FE estimate can vary from sample to sample, even if the marginal effects are the same. Consequently, caution is called for when comparing linear FE results across studies. For example, if the relationship between family income and child outcomes is approximately concave, a linear FE estimate will be declining with the share of within-family income variation that is experienced by rich families, which might be quite different across countries or subsamples.

E. *Decompositions with Heterogeneous Marginal Effects*

To focus attention on the implications of nonlinearity when the estimated model is assumed to be linear, both model (1) and (2) restricted the functional form to be the same for all children, that is $\gamma_{ci} = \gamma_c$ for all i. As shown in the Appendix, with heterogeneity in the marginal effects, that is $\gamma_{ci} \neq \gamma_{c'}$ for $i \neq i'$, it follows straightforwardly that

$$\beta(OLS) = \sum_{c=1}^{\bar{c}} \gamma_c(OLS)\, w_c(OLS), \tag{9}$$

$$\beta(Z) = \sum_{c=1}^{\bar{c}} w_c(Z)\,\gamma_c(Z), \tag{10}$$

$$\beta(FE) = \sum_{c=1}^{\bar{c}} w_c(FE)\,\gamma_c(FE), \tag{11}$$

[10]Note that the decomposition of the sample analog of $\beta(FE)$ has exactly the same form: $\hat{\beta}(FE) = \sum_{c=1}^{\bar{c}} \hat{w}_c(FE)\hat{\gamma}_c(FE)$, where $\hat{\beta}(FE)$, $\hat{w}_c(FE)$, and $\hat{\gamma}_c(FE)$ are the sample analogs of the estimands.

where $\gamma_c(OLS)$, $\gamma_c(Z)$, and $\gamma_c(FE)$ are the OLS, IV, and FE estimand of the cth marginal effect, respectively. With heterogeneity in marginal effects, these estimands will generally differ even in the absence of endogeneity bias because they identify the *average* marginal effects for different subgroups. For example, the IV estimand for the cth marginal effect identifes the marginal local average treatment effect (Angrist and Imbens 1995).

F. *Comparison of Linear Estimators*

As is evident from equations (9)–(11), with heterogeneity in marginal effects, the linear OLS, IV, and FE estimators will identify different weighted averages of different marginal effects estimates. We now provide pairwise comparisons that enable us to partition the differences in the linear estimands into different weights and different marginal effects estimates. Our approach is analogous to the Blinder-Oaxaca decomposition, in that differences in mean outcomes of two groups can be partitioned into both differing characteristics and differing effects between the two groups. Note, however, that our decomposition is path independent, that is, the decomposition results are independent of the order in which the decomposition is performed.

We begin by decomposing the difference between the linear FE estimand and the linear IV estimand. Let

$$\beta(FE,Z) = \sum_{c=1}^{\bar{c}} w_c(Z)\gamma_c(FE)$$

define the IV weighted FE estimand, and

$$\beta(Z,FE) = \sum_{c=1}^{\bar{c}} w_c(FE)\gamma_c(Z)$$

define the FE weighted IV estimand. Then, we can write

$$\begin{aligned}(12)\quad \beta(FE) - \beta(Z) &= \frac{1}{2}(\{\beta(FE) - \beta(FE,Z)\} + \{\beta(FE,Z) - \beta(Z)\}) \\ &\quad + \frac{1}{2}(\{\beta(FE) - \beta(Z,FE)\} + \{\beta(Z,FE) - \beta(Z)\}) \\ &= \underbrace{\frac{1}{2}\left\{\sum_{c=1}^{\bar{c}}[w_c(FE) - w_c(Z)][\gamma_c(Z) + \gamma_c(FE)]\right\}}_{\text{weights}} \\ &\quad + \underbrace{\frac{1}{2}\left\{\sum_{c=1}^{\bar{c}}[\gamma_c(FE) - \gamma_c(Z)][w_c(FE) + w_c(Z)]\right\}}_{\text{marginal effects}}.\end{aligned}$$

This expression shows that $\beta(FE)$ will, in general, differ from $\beta(Z)$, even when both the IV assumptions and the FE assumptions are satisfied. The first term, denoted *weights*, tells us what would have been the difference between $\beta(FE)$ and $\beta(Z)$ if the marginal effect estimates were the same ($\gamma_c(FE) = \gamma_c(Z)$) and only the distribution of weights differed; that is, the contribution of differences in weighting to the difference in the linear FE and IV estimates. The second term, denoted *marginal effects*, tells us what would have been the difference between $\beta(FE)$ and $\beta(Z)$ if the distributions of weights were the same ($w_c(FE) = w_c(Z)$) and only the marginal effects estimates differed; that is, the contribution of differences in marginal effects to the difference in the linear FE and IV estimates.

Using the same procedure, it follows straightforwardly that

$$(13)\quad \beta(OLS) - \beta(FE) = \underbrace{\frac{1}{2}\left\{\sum_{c=1}^{\bar{c}} [w_c(OLS) - w_c(FE)][\gamma_c(FE) + \gamma_c(OLS)]\right\}}_{\text{weights}} + \underbrace{\frac{1}{2}\left\{\sum_{c=1}^{\bar{c}} [\gamma_c(OLS) - \gamma_c(FE)][w_c(FE) + w_c(OLS)]\right\}}_{\text{marginal effects}}$$

decomposes the difference between the linear OLS estimand and the linear FE estimand, and

$$(14)\quad \beta(OLS) - \beta(Z) = \underbrace{\frac{1}{2}\left\{\sum_{c=1}^{\bar{c}} [w_c(OLS) - w_c(Z)][\gamma_c(Z) + \gamma_c(OLS)]\right\}}_{\text{weights}} + \underbrace{\frac{1}{2}\left\{\sum_{c=1}^{\bar{c}} [\gamma_c(OLS) - \gamma_c(Z)][w_c(Z) + w_c(OLS)]\right\}}_{\text{marginal effects}}$$

decomposes the difference between the linear OLS estimand and the linear IV estimand.

Equations (13) and (14) show that $\beta(OLS)$ will, in general, differ from $\beta(FE)$ and $\beta(Z)$, even when family income is exogenously determined, $\Delta_c = 0$ for all c. Both decompositions consist of two terms. The first term captures the difference due to differing weights, and the second term captures the difference due to differing marginal effects estimates. Note that the marginal effects term in (13) and (14) may be nonzero not only because of heterogeneous marginal effects, as in (12), but also because of endogeneity bias.

II. Data and Background

This section describes our data and discusses the natural experiment used as an instrument for family income before displaying descriptive statistics.

A. *Data and Sample Selection*

As in Løken (2010), our empirical analysis uses a rich longitudinal dataset containing records for every Norwegian from 1967 to 2006. The variables captured in this dataset include individual demographic information (sex, birth year, marital status, number of children, etc.), and socioeconomic data (years of education, IQ, income, etc). Importantly, the dataset includes personal identifiers for one's parents, allowing us to link children to their parents and siblings, as well as family identifiers, allowing us to link spouses. Moreover, the dataset includes geographic identifiers for county of birth.

In the empirical analysis, we use two samples. As explained below, in the IV estimation, we select children born in the treatment and control counties in the years 1965, 1967, 1968, and 1969. We also perform FE estimation on a sample of sibling pairs. To get sufficient precision in the FE estimates, our FE sample comprises sibling pairs born in 1965, and between 1967 and 1977 in the treatment and control counties. This serves two purposes. First, by increasing the sample size, we get an adequate number of siblings. Second, by having siblings further spaced apart, we obtain more within-family income variation. To provide direct comparison of the OLS versus IV results, as well as the OLS versus FE results, we perform OLS estimation in both the IV and FE sample.

Throughout the paper, we use three different measures for children's outcomes: years of education, whether the individual is a high school dropout, and an IQ test score. Years of education is defined as the number of completed years of education in 2006, whereas high school dropout is defined as not obtaining a three year high school diploma by 2006. In 2006, the children in our sample are at least 29 years old, which ensures that almost all have completed their education. Unlike these two outcome measures, the IQ test score is only available for males because they are collected from military records, and military service is compulsory for men only. Before entering the military, their medical and psychological suitability is assessed; this occurs for the great majority between their eighteenth and twentieth birthdays. The IQ test score at these ages is particularly interesting as it is about the time of entry to the labor market or to higher education. The IQ test score is a composite score from three timed tests—arithmetic, word similarities, and figures (see Thrane 1977, and Sundet et al. 2004, 2005). The composite IQ test score is an unweighted mean of the three subtests. The IQ score is reported in *stanine* (Standard Nine) units, a method of standardizing raw scores into a nine point standard scale that has a discrete approximation to a normal distribution, a mean of five, and a standard deviation of two. We have IQ scores on about 84 percent of all Norwegian men born in the years we consider.[11]

We follow Løken (2010) in our definition of family income. Income is taken from tax registers, and includes all market income, from wages and self-employment, as

[11] Eide et al. (2005) examine patterns of missing IQ data for the men in the 1967–1987 cohorts. Of those, 1.2 percent died within 1 year and 0.9 percent died between 1 year of age and registering with the military at about age 18. About 1 percent of men had emigrated before age 18, and 1.4 percent were exempted because they were permanently disabled. An additional 6.2 percent of scores are missing for a variety of reasons, most notably foreign citizenship.

well as (taxable) cash benefits, such as unemployment benefits, disability benefits, and sickness pay. We deflated the income to real 1999 income, by using the average yearly consumer price index. In every year, we add the income of the child's mother to her spouse's income (if she is married) to create one variable capturing annual family income. This means that we measure family income as the total income of the family that the child lived in, regardless of whether the spouse of the mother is the child's biological father. We then define family income during childhood as the average annual family income from age 1 until age 11.

B. *Natural Experiment*

We follow Løken (2010) in exploiting time and regional variation in the economic boom that followed the initial (offshore) discovery of oil in Norway as a source to exogenous variation in childhood family income.[12] Our motivation for using the initial discovery of oil as a natural experiment is twofold. First, the extent to which childhood family income of children born in a given year is affected by the subsequent oil boom depends on the geographical proximity of their place of birth to the offshore oil fields. And second, for children born in the same place, the effect of the oil boom on childhood family income depended on their year of birth. In particular, our first stage of the 2SLS will be a difference-in-differences specification, exploiting that the oil boom most strongly affected the childhood family income of children born in the years right before the discovery of oil, in the county located just off the coast of the offshore oil fields.

At the end of 1969, the first major oil discovery was made in the North Sea in Norway, and in June 1970, the public was informed of it. Eventually, the discovery of oil fueled the entire Norwegian economy, but Rogaland county was the first and most strongly influenced because the main oil production in the relevant period was located off the coast there. The large increase in labor demand from the oil industry gradually spilled over into higher wages also for other types of jobs.

To avoid threats to the validity of the instrument from endogenous migration, the sample used in the IV estimation consists of the cohorts born just prior to the initial discovery of oil in Norway. The *treatment group* consists of the subsample of children born in Rogaland county, whose families were exposed to the oil boom to a greater extent than families elsewhere in Norway because of Rogaland's geographical proximity to the offshore oil fields. As these children grew up, this led to a rise in family income in Rogaland compared to other counties of Norway. The *control group* comprises children born in ten other counties that are geographically distant from the offshore oil fields, but with similar family and child characteristics as those in Rogaland.[13] In general, there are long driving distances between the populated

[12] See Løken (2010) for a detailed discussion of the oil boom, and a number of results supporting the validity of this natural experiment as a source of exogenous variation in family income.

[13] These counties are Sør-Trøndelag, Hedmark, Vestfold, Aust-Agder, Oppland, Telemark, Sogn og Fjordande, Møre og Romsdal, Nord-Trøndelag, and Buskerud. The eight excluded counties are: Oslo and Akershus, comprising the capital and the surrounding urban area; Finmark, Troms, and Nordland, the three northernmost counties; and Aust-Agder and Hordaland, the close neighboring counties to Rogaland.

areas of the counties of Norway, as they are mostly far apart or partitioned by mountains and/or the fjord-gashed shoreline.

As discussed in Løken (2010), before Norway discovered oil, Rogaland was a typical Norwegian county whose main economic activities revolved around fish and agriculture. This is mirrored in the descriptive statistics reported in Table 2, showing that children in Rogaland had very similar individual and parental characteristics as those from the control counties. It also should be noted that the oil boom had little, if any, impact on local public spending on schooling in Rogaland compared to other Norwegian counties. This is in part because of the unitary and federally funded school system in Norway, but also due to the fact that public oil revenues went directly to the central government and were redistributed to the counties independently of proximity to the offshore oil fields. Hence, the instrument will pick up variation in family income due to higher labor demand in the affected county, rather than greater public goods expenditures in this area. When the children in our sample were old enough to start their post-secondary education, the oil boom was already incorporated in the Norwegian economy, so that there were no (observable) differences between Rogaland and the rest of Norway in returns to education.

Even if children from the treatment group have very similar observable characteristics as children from the control group, we cannot rule out that they have different unobserved family and child characteristics, and therefore would have different educational attainment and test scores in the absence of the discovery of oil. To address this concern, we not only include children from the treatment and control group born in the years immediately before the reform, 1967–1969, but also children born a couple of years earlier, in 1965. Our instrument is defined as belonging to the treatment group interacted with being born in the years 1967–1969. The first stage is then a difference-in-differences estimate of the effect on family income of being born in Rogaland instead of one of the control counties for the 1967–1969 cohorts compared to the 1965 cohort. Our first and second stage specifications therefore include fixed effects for birth cohort and county of birth, controlling for unobserved permanent differences between children born in different years as well as between children born in different areas. Our estimates are very similar when excluding children born in 1965, as Løken (2010) does, in which case we cannot control for unobserved differences between children born in different areas.

C. *Descriptive Statistics*

Table 1 provides descriptive statistics for the FE and IV samples. As displayed in the first row, the IV sample consists of more than 120,000 children. As expected, more fathers than mothers have attended college, and fathers are, on average, a few years older than mothers. For our outcome variables, we see that average education in the sample is 12.4 years and about 30 percent of the sample have not obtained a high school diploma. The average IQ test score of boys are 5 out of a scale of 1–9. Finally, we see that average childhood family income is around NOK 252,000 (USD $43,450).

As shown in Table 1, the FE sample consists of more than 202,000 children. We see that the IV and FE sample are quite similar in terms of observable characteristics.

Table 1—Descriptive Statistics for IV and FE Sample

	IV sample		FE sample	
	Mean	SD	Mean	SD
Female	0.49	0.50	0.49	0.50
Number of siblings	2.13	1.29	1.98	1.11
Birth order	2.20	1.28	1.91	1.00
Mother college	0.07	0.26	0.09	0.28
Father college	0.16	0.37	0.17	0.38
Mother's age when child is born	26.7	5.98	25.2	4.77
Father's age when child is born	29.9	7.90	28.3	5.81
Education in 2006	12.39	2.50	12.60	2.50
Dropout rate from high school	0.30	0.46	0.26	0.44
Adult IQ (boys only)	5.01	1.82	5.11	1.80
Family income (child aged 2–12) in NOK 10,000	25.2	10.1	27.5	10.1
Height (males only)	179.7	6.5	179.8	6.6
Observations	121,122		202,424	

Notes: The IV sample consists of children born in 1965 and 1967–1979 in the following counties: Rogaland, Sør-Trøndelag, Hedmark, Vestfold, Aust-Agder, Oppland, Telemark, Sogn og Fjordande, Møre og Romsdal, Nord-Trøndelag, and Buskerud. The FE sample consists of children born in 1965 and 1967–1977 from the same counties.

As we have added younger cohorts to the sample, the children in the FE sample are, on average, younger than those in the IV sample. We also see that they have slightly fewer siblings, which is attributable to the declining fertility trend over time. Parents are younger at the time of birth and have slightly more completed education. We also see that children in the FE sample are performing better in terms of the outcome measures, most likely a result of the increasing trend in educational attainment across cohorts.

Table 2 shows differences in the average outcomes for children from the treatment and control group who were born in 1967–1969, as well as those born in 1965. As is evident from the table, children from the treatment group (who were born in Rogaland) have somewhat lower educational attainment and adult IQ, as well as slightly higher dropout rates, compared to children from the control group (who were born in the other counties). We also see that these differences change very little across the cohorts. In a linear IV framework, this is suggestive of a small, if any, effect of family income on child outcomes when using the oil boom as the instrument. This is because the reduced form of the 2SLS (without controls for child and family characteristics) would be equal to the mean difference in the outcome of interest between the treatment and control group for the 1967–1969 cohorts, subtracted from the same mean difference for the 1965 cohort (i.e., a difference-in-differences estimator).

Table 2 also displays the mean differences in characteristics and family income for children from the treatment and control group who were born in 1967–1969, as well as those born in 1965. We immediately see that the treatment and control groups are quite similar in terms of observable characteristics, and moreover, that these differences change little across cohorts. In contrast, average family income is substantially higher in the treatment group compared to the control group, especially for the 1967–1969 cohorts, which forms the basis for our first stage regressions.

Table 2—Mean Differences between Children from the Treatment and Control Group, by Birth Cohort

	Levels treatment 1967–1969	Difference (SE) treatment-control 1967–1969	Difference (SE) treatment-control 1965
Female	0.49	−0.001 (0.005)	−0.003 (0.008)
Number of siblings	2.08	0.23 (0.011)	0.19 (0.021)
Birth order	2.18	0.11 (0.011)	0.05 (0.021)
Mother college	0.07	−0.008 (0.002)	−0.013 (0.004)
Father college	0.16	−0.006 (0.003)	0.003 (0.006)
Mother's age when child is born	26.5	0.22 (0.053)	0.09 (0.101)
Father's age when child is born	29.5	−0.14 (0.069)	−0.36 (0.138)
Education in 2006	12.27	−0.201 (0.023)	−0.209 (0.040)
Dropout rate from high school	0.29	0.004 (0.004)	0.012 (0.008)
Adult IQ (males only)	5.06	0.017 (0.024)	0.049 (0.043)
Family income (child aged 2–12) in NOK 10,000	26.3	2.65 (0.090)	1.87 (0.157)
Height (males only)	179.3	−0.58 (0.083)	−0.51 (0.144)
Observations	14,759	91,164	29,958

Notes: The treatment group consists of children born in Rogaland. The control group consists of children born in Sør-Trøndelag, Hedmark, Vestfold, Aust-Agder, Oppland, Telemark, Sogn og Fjordande, Møre og Romsdal, Nord-Trøndelag, and Buskerud.

III. Empirical Results

This section outlines our empirical models used in the OLS, IV, and FE estimation before discussing our empirical results.

A. *Empirical Models*

The main empirical model used in the literature is specified as

$$\text{Model 1: } y_i = \mu_0 + \beta C_i + X_i'\delta + \epsilon_i,$$

where y_i is some outcome, C_i is childhood family income, and X_i is a set of controls. Throughout our paper, X_i includes fixed effects for birth cohort and county of birth, as well as dummy variables for child's birth order, number of siblings, gender, and parent's age and college attendance. All the control variables are measured in the year the child is born, and therefore before our variable of interest C_i.

To allow the marginal effects of additional income to vary across the family income distribution, we use the following model:

Model 2: $y_i = \mu_0 + \beta_1 C_i + \beta_2 C_i^2 + X_i'\delta + \upsilon_i.$

This quadratic specification in family income conforms to theory suggesting a concave relationship, and aims at achieving a reasonable tradeoff between a flexible functional form in family income and precision in the IV and FE estimation. The robustness analysis examines the sensitivity of the results produced by Model 2 to the inclusion of higher order polynomials in family income.

To perform 2SLS estimation of Models 1 and 2, we use the following first stage specifications, where Z_i is a set of instruments and X_i is the same set of controls as above:

First stage, Model 1:

$$C_i = Z_i'\lambda + X_i'\rho + \eta_i;$$

First stages, Model 2:

$$C_i = Z_i'\lambda + X_i'\rho + \eta_i \quad \text{(First Stage I)}$$

$$C_i^2 = Z_i'\theta + X_i'\xi + \nu_i \quad \text{(First Stage II).}$$

When performing IV estimation of Model 1, we first use a single binary instrument, equal to one if the child is born in Rogaland in the years 1967–1969, and zero otherwise. We refer to this instrument as the "Rogaland dummy variable." However, to identify the parameters of both the linear and the squared family income terms in Model 2, we need more than one instrument. To construct multiple instruments, we use two different strategies. Both strategies exploit the fact that if the Rogaland dummy variable is a valid instrument, then under an assumption that ϵ_i is mean-independent of the included covariates X_i (a necessary assumption for consistent IV estimation of Model 1), any function of the Rogaland dummy variable and the X_i are valid instruments.

The first IV strategy interacts the Rogaland dummy variable with some of the covariates, and uses this set of instruments in the first stage specifications (First Stage I and II) of Model 2. Our reason for not interacting the Rogaland dummy variable with all the included control variables is that such a procedure would introduce a large number of overidentifying restrictions, which could increase the small sample bias of the IV estimator (see e.g., Staiger and Stock 1997). As a tradeoff between small sample bias and efficiency in the IV estimation, our main specification interacts the Rogaland dummy variable with five control variables: indicator for father attended college, indicator for mother attended college, father's age, mother's age, and an indicator for large family size (three or more siblings). The reason for choosing these control variables is that they generate the strongest first stage results. Importantly, to provide a direct comparison between the IV results of Models 1 and 2, we will report 2SLS results from both models using the same set of interaction instruments. As a robustness check, we also report IV results using only a subset of these interactions, as well as the IV results from interacting the Rogaland dummy variable with every included covariate.[14]

The second IV strategy uses predicted family income and predicted family income squared as the instruments. This strategy follows closely the IV literature where the predicted treatment is used as the instrument in a conventional 2SLS procedure.[15] The predicted family income instruments are constructed by regressing family income on the controls, the Rogaland dummy variable, and the five interactions discussed above. From these regression coefficients, we predict family income for each child. Finally, we apply the standard 2SLS procedure using predicted family income and predicted family income squared as instruments, controlling for the $\mathbf{X}_i$ variables. To provide a direct comparison between the IV results of Models 1 and 2, we also report 2SLS results from both models using the same set of predicted family income instruments.

B. *Linear and Quadratic Estimates*

Columns 1–3 of Table 3 report the linear OLS, IV, and FE estimates for our three outcome measures: years of education, high school dropout, and adult IQ. The strong first stage results are reported in Appendix Table B1. From panel A of Table 3, we see that our precise OLS result indicates a positive and sizable association between family income and children's educational attainment and IQ as adults. To get a perspective on the magnitude of the parameter estimates, a standard deviation increase in family income (NOK 101,000 or USD $17,414) is associated with an increase in years of schooling of about 0.4, a fall in high school dropout rates in the range of 5–6 percentage points, and a rise in the IQ test score of more than 0.125 of a standard deviation. Comparing the results in panels A and C, we see that the linear OLS estimates are quite similar in the IV and FE sample.[16]

In panels B and D of Table 3, we report linear IV and FE results. The FE estimates of family income on children's IQ and educational attainment are close to zero and significantly different from the linear OLS estimates. In the spirit of previous studies, these results would be interpreted as suggesting little, if any, causal effect of family income on children's outcomes, as well as significant endogeneity bias in the OLS estimates. The IV estimates also show no sign of significant effects of family income on child outcome, although they are too imprecisely estimated to rule out some effect.

[14] Note also that assuming the Model 1 specification as the data generating process for C_i would necessarily imply that interactions of covariates and the Rogaland dummy variable should be included in the First Stage II specification for C_i^2.

[15] Under the assumption that the included covariates and the instruments are mean-independent of the regression error, $E[v_i | \mathbf{X}_i, Z_i] = 0$, we can use any function of the $\mathbf{X}$ and Z variables to form instruments. We use $E[C_i | \mathbf{X}_i, Z_i]$ as the predicted value of childhood income given the covariates and instruments. To the extent that these predicted family income instruments are more highly correlated with the endogenous level of income, they generate efficiency gains. Wooldridge (2002), Carneiro, Heckman, and Vytlacil (2003), and Mogstad and Wiswall (2011b) provide examples of analysis using the predicted treatment as the instrument. In these applications, they find a substantial improvement in the precision of the IV estimates using the predicted treatment instruments over the IV estimates using the instruments directly.

[16] When considering the adult IQ results, Table 3 displays a larger reduction in sample size in the FE estimation compared to the IV estimation. This is because the FE estimates are identified from within-family income variation among brothers.

Table 3—Linear OLS, IV, and FE Estimates

	Education	Dropout	Adult IQ (males only)
Panel A. Linear OLS			
Family income in NOK 10,000	0.043***	−0.006***	0.028***
	(0.001)	(0.000)	(0.001)
Panel B. Linear IV			
Instrument: born in Rogaland in 67–69			
Family income in NOK 10,000	0.022	−0.012	0.033
	(0.057)	(0.011)	(0.023)
Observations	121,122	121,122	57,788
Panel C. Linear OLS			
Family income in NOK 10,000	0.041***	−0.005***	0.024***
	(0.001)	(0.000)	(0.001)
Panel D. Linear FE			
Family income in NOK 10,000	0.000	−0.001	−0.001
	(0.003)	(0.001)	(0.004)
Observations	202,424	202,424	55,866

Notes: This table reports OLS, IV, and FE estimates of Model 1. Panels A and B use the IV sample, whereas panels C and D use the FE sample. Panel B uses the born in Rogaland in 1967–1969 dummy variable as the only instrument. A full set of controls is used in all regressions. Standard errors in parentheses are heteroskedastic robust.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 4 reports OLS, IV, and FE results from Model 2, where we have relaxed the linearity restriction in Model 1 by including a squared term in family income. The strong first stage results are reported in Appendix Table B1. Panel B of Table 4 reports IV results using the interacted instruments, and panel C shows the IV results applying the predicted family income instruments. Our main finding is that there is an increasing, concave relationship between family income and children's outcomes, with large marginal effects in the lower parts of the family income distribution. This holds true for both the OLS and FE results, as well as in the IV estimates, regardless of the choice of instruments.

Figure 1 graphs the predicted effects of family income across the family income distribution. We can see that the predicted effects from the FE estimates are smaller and display a less concave pattern than the IV estimates. To be concrete, the FE estimates suggest that a standard deviation increase in income (NOK 101,000, USD $17,414) produces 0.22 additional years of education for a child from a family with income of NOK 150,000, whereas a child from a family with income of NOK 300,000 only achieves an extra 0.02 years of education. In comparison, the linear IV estimates using the predicted family income instruments indicates that such an increase in family income would generate 0.74 additional years of education for a child from the poor family, whereas the child from the richer family would gain as little as 0.05 years of education. Below, we discuss the likely reasons for the disparity in the predicted marginal effects.

Table 4—Quadratic OLS, IV, and FE Estimates

	Dependent variables		
	Education	Dropout	Adult IQ (males only)
Panel A. Quadratic OLS			
Family income in NOK 10,000	0.051***	−0.010***	0.032***
	(0.002)	(0.000)	(0.002)
Quadratic income (× 100)	−0.016***	0.008***	−0.008*
	(0.005)	(0.001)	(0.004)
Panel B. Quadratic IV			
Instruments: Interactions			
Family income in NOK 10,000	0.180**	−0.030*	0.234**
	(0.087)	(0.016)	(0.109)
Quadratic income (× 100)	−0.302*	0.042*	−0.401**
	(0.164)	(0.021)	(0.210)
Panel C. Quadratic IV			
Instruments: Predicted income			
Family income in NOK 10,000	0.142**	−0.057***	0.195***
	(0.072)	(0.013)	(0.070)
Quadratic income (× 100)	−0.228**	0.097***	−0.323***
	(0.107)	(0.019)	(0.106)
Observations	121,122	121,122	57,788
Panel D. Quadratic OLS			
Family income in NOK 10,000	0.069***	−0.012***	0.033***
	(0.002)	(0.000)	(0.003)
Quadratic income (× 100)	−0.050***	0.013***	−0.017***
	(0.003)	(0.001)	(0.004)
Panel E. Quadratic FE			
Family income in NOK 10,000	0.041***	−0.006***	0.019*
	(0.008)	(0.002)	(0.010)
Quadratic income (× 100)	−0.065***	0.008***	−0.031**
	(0.013)	(0.002)	(0.014)
Observations	202,424	202,424	55,866

Notes: This table reports OLS, IV, and FE estimates of Model 2. Panels A, B, and C use the IV sample, whereas panels D and E use the FE sample. Panel B uses the set of interacted instruments (born in Rogaland in 1967–1969 dummy variable and interacting the born in Rogaland in 1967–1969 dummy variable with father's college, mother's college, father's age, mother's age, and large family) in First Stages I and II, whereas panel C uses the predicted family income instruments based on the same set of interacted instruments in First Stages I and II. A full set of controls is used in all regressions. Note that the quadratic income regressor is linear income (in NOK 10,000) multiplied by 100, hence quadratic income is in NOK 1,000,000. Standard errors in parentheses are heteroskedastic robust.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

C. *OLS, IV, and FE Weights*

To understand why the linear FE and IV estimates miss the substantial impact of family income on child outcomes, we compute the set of weights for each estimator based on the decompositions in equations (9)–(11). Figure 2 graphs the distribution of weights for the linear OLS, IV, and FE estimates. In order to compute the weights, we have discretized the income distribution using family income margins of NOK 10,000

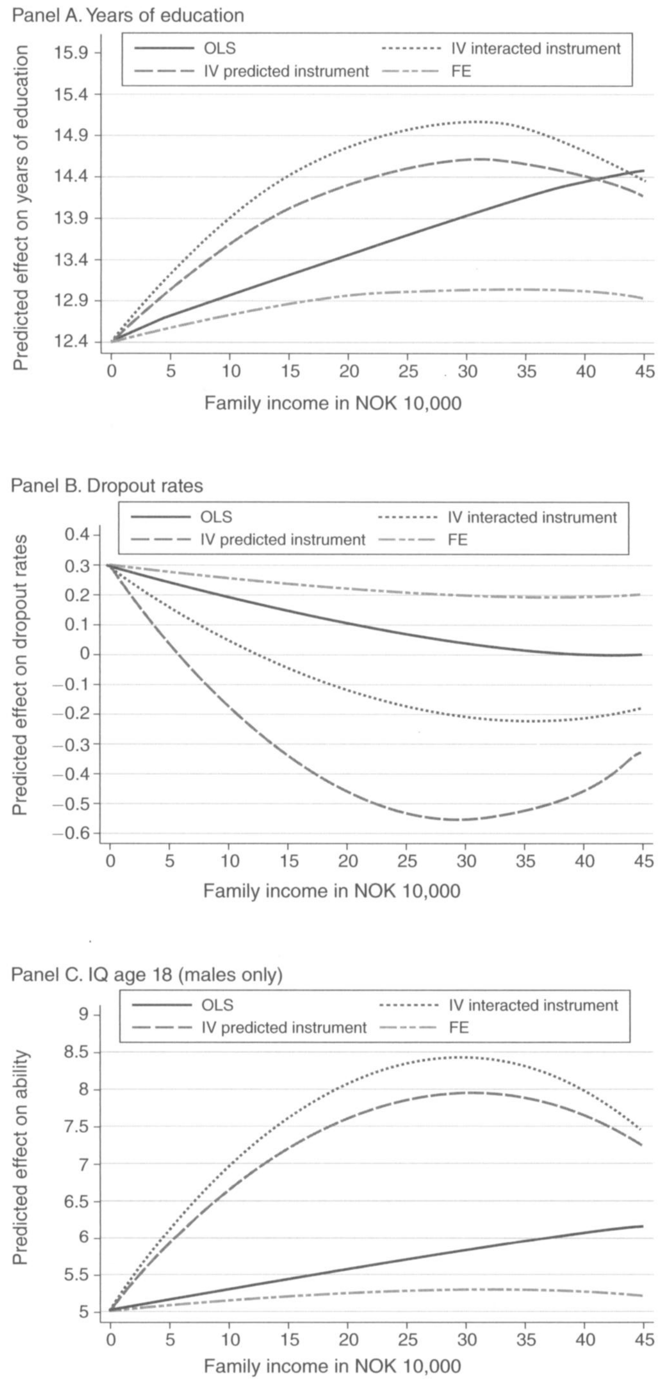

Figure 1. Predicted Effects from Quadratic Model

Notes: This figure shows the predicted (total) effects based on the estimates of Model 2. Each graph shows predicted effects from OLS estimates (panel A of Table 4), FE estimates (panel E of Table 4), IV estimates with interacted instruments (panel B of Table 4), and IV estimates with predicted family income instruments (panel C of Table 4).

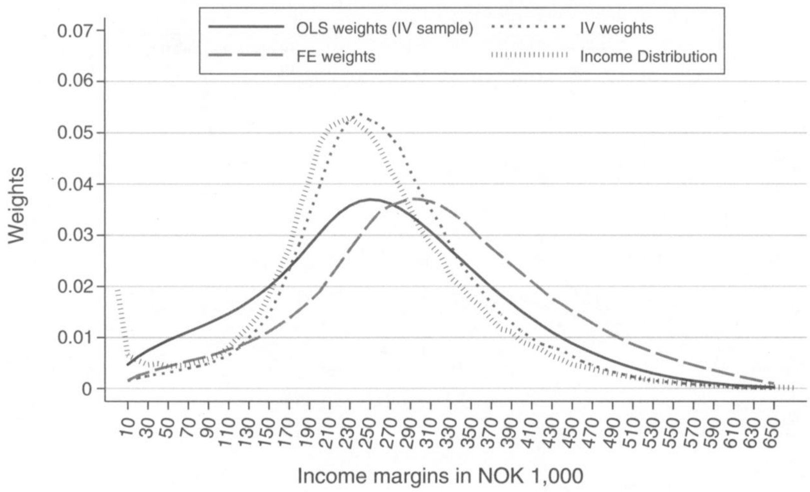

Figure 2. Linear OLS, IV, and FE Weights

Notes: This figure reports the weights for the linear OLS and IV estimates (using the IV sample) shown in panels A and B, and the linear FE estimate (using the FE sample) shown in panel D of Table 3. To compute these weights, we use the decomposition in (9)–(11), and income margins of NOK 10,000. For comparison, this figure also graphs the family income distribution.

(USD \$1,667). The figure reveals that the linear IV estimate assigns relatively little weight to the large marginal effects in the lower part of the family income distribution, reflecting that the oil boom did not do that much for the family income of poor families. It is also evident that well-off families in Norway experience most within-family income variation, implying that the relatively small marginal effects in the middle and upper part of the family income distribution contribute the most to the linear FE estimate. In comparison, the linear OLS estimator weights the different margins more evenly, assigning several times more weight to marginal effects in the lower part of the family income distribution than the linear IV and FE estimator.

D. *Comparison of Linear Estimates*

To examine how much the weights explain the differences across the linear estimates, we use the decomposition approach given by equations (12)–(14). Table 5 displays average marginal effects and decomposition results for years of schooling. The average marginal effects are based on the estimates of Model 2, reported in Table 4. Each cell in panel A shows a weighted sum of marginal effect estimates. The first (second or third) row weights the OLS (IV or FE) estimates of the marginal effects with OLS weights (column 2), IV weights (column 3), and FE weights (column 4). Panel B shows the decomposition results. Column 1 shows the absolute difference in the linear estimates, whereas columns 2 and 3 report the contribution of differences in weights and marginal effects, respectively.

With homogeneity in the marginal effects (and no measurement error), $\beta(Z)$ and $\beta(FE)$ should differ only because of different weighting. We expect, however, the OLS estimates to be (upward) biased, in which case $\beta(OLS)$ and $\beta(Z)$ or $\beta(OLS)$

TABLE 5—AVERAGE MARGINAL EFFECTS AND DECOMPOSITION RESULTS

	Weights		
	OLS	IV	FE
Panel A. Estimates			
OLS	0.0429	0.0428	0.0412
IV	0.0249	0.0222	−0.0068
FE	0.0093	0.0087	0.0019
	Decomposition		
	Absolute difference	Weights	Marginal effects
Panel B. Estimates			
$\beta(OLS) - \beta(Z)$	0.0207	7%	93%
$\beta(OLS) - \beta(FE)$	0.0410	11%	89%
$\beta(Z) - \beta(FE)$	0.0203	88%	12%

Notes: Panel A reports average marginal effects and panel B reports decomposition results. Panel A: Each cell shows a weighted sum of marginal effect estimates. The first (second or third) row weights the OLS (IV or FE) estimates of the marginal effects with OLS weights (column 1), IV weights (column 2), and FE weights (column 3). The marginal effects are based on the estimates of Model 2, reported in Table 4. The IV estimation uses interacted instruments. The weights are based on equations (9)–(11). To compute the weights, we use income margins of NOK 10,000. Panel B: column 1 shows the absolute difference between the linear estimates, whereas columns 2 and 3 report the contribution of differences in weights and marginal effects, respectively. The decomposition results are based on equations (12)–(14), with average marginal effects from panel A.

and $\beta(FE)$ may differ, also due to different marginal effect estimates. Table 5 shows that for the comparison between $\beta(FE)$ and $\beta(Z)$, the choice of weights is what really matters. While only 12 percent of the difference between the estimates of $\beta(Z)$ and $\beta(FE)$ is attributable to different marginal effects, the contribution of different weighting is as large as 88 percent. In comparison, different marginal effect estimates is clearly the driving factor behind the relatively high estimate of $\beta(OLS)$.

Table 6 summarizes the decomposition results for all the outcomes. Given the bias in the OLS estimates making them less informative, this table focuses attention on the differences between $\beta(Z)$ and $\beta(FE)$. Panel A reports results based on the interaction instruments, whereas panel B reports results based on the predicted instruments.

Column 1 demonstrates that the IV estimates exceed (in absolute value) the FE estimates, as found in previous studies. Dahl and Lochner (2011) discuss several possible explanations. One is that income is also noisily measured, so that the FE estimates suffer from more attenuation bias than the IV estimates. It is also possible that the effect of family income is greater for the complier group to our instruments, than for other children. Another possibility is that endogenous income shocks are creating downward bias in the FE estimates. It could also be that transitory income shocks to family income creating within-family income variation may, at least partly, be smoothed out by intertemporal income transfers, lowering the estimated effect of family income in the FE estimates. All these explanations pertain to differences in particular marginal effects estimates, such that $\gamma_c(FE) > \gamma_c(Z)$. In contrast, much less attention has been devoted to the possibility that the IV estimate exceeds the FE estimate because the latter assigns less weight to the sizable marginal effects in the lower part of the distribution.

TABLE 6—DECOMPOSITION RESULTS: LINEAR IV AND FE ESTIMATES

		Decomposition	
Outcome	$\beta(Z) - \beta(FE)$	Weights	Marginal effects
Panel A. Instruments: interactions			
Education	0.0203	88%	12%
Dropout	−0.0068	35%	65%
Adult IQ	0.0253	82%	18%
		Decomposition	
	$\beta(Z) - \beta(FE)$	Weights	Marginal effects
Panel B. Instruments: predicted			
Education	0.0213	67%	33%
Dropout	−0.0062	81%	19%
Adult IQ	0.0269	64%	36%

Notes: This table reports decomposition results for linear IV and FE estimates, with IV estimates based on interacted instruments (panel A) and predicted instruments (panel B). Column 1 shows the absolute difference between the linear estimates, whereas columns 2 and 3 report the contribution of differences in weights and marginal effects, respectively. The decomposition results are based on equations (12)–(14), with average marginal effects from panel A of Table 5.

Columns 2 and 3 explore the contribution of differences in weights and marginal effects to the discrepancy between $\beta(Z)$ and $\beta(FE)$. Our main finding is that weighting matters a lot. The only case where marginal effects matter the most is for dropout rates based on the interaction instruments. It should be noted, however, that unlike the other quadratic IV results, the point estimates for dropout rates based on interacted instruments are quite imprecise, in which case sampling error may generate substantial difference between the estimates of $\beta(Z)$ and $\beta(FE)$.

E. *Robustness Analysis*

This subsection discusses a number of robustness checks, supporting the validity of our main results.

Functional Form.—Our quadratic specification is intended to achieve a reasonable tradeoff between flexibility in functional form and achieving sufficient precision. We have, however, also estimated the effect of family income with a cubic term in family income. The coefficients associated with the cubic term are insignificant, indicating that the concave specification might be reasonable. However, we admittedly cannot rule out that an even more flexible specification would provide a better approximation of the causal relationship between family income and child outcome. At the very least, our quadratic model nests the linear model, and is therefore an improvement over the linear specification in family income used in previous studies.[17]

[17] Rather than impose a particular nonlinear functional form, one could use a more flexible nonparametric approach to estimation (e.g., the partially linear model of Robinson 1988). In the OLS estimation, we have attempted more flexible functional forms (e.g., including using a series of 65 dummy variables for income margins of NOK 10,000), confirming a concave association between family income and child outcomes. However, while these approaches are feasible for OLS estimation, nonparametric IV methods tend to suffer from the "ill posed inverse" problem making such estimation difficult in practice (see Horowitz 2011and Newey and Powell 2003).

Choice of Instruments.—With heterogeneity in marginal effects, IV estimates will generally vary with the choice of instruments because they identify the marginal effects of different subgroups. Indeed, the quadratic IV estimates reported in Table 4 differ somewhat, though not significantly, depending on whether we use the interacted instruments or the predicted family income instruments. A concern is therefore that the use of different instruments explain the discrepancy between the IV estimates from the linear and quadratic model.

Table B2 addresses this concern, reporting IV results from Model 1 based on the same instruments as used in the IV estimates of Model 2. Comparing the results from the linear and quadratic specification using the *same* set of instruments, it is clear the role of the linearity restriction in masking the family income effects. The linear IV estimates are always insignificantly different from zero, whereas the quadratic IV estimates using the same set of instruments show large and significant effects of family income. Hence, we can conclude for a given set of instruments, inappropriately assuming a linear model is important for the conclusions about the effects of family income on child outcome.

Small Sample Bias.—The small sample properties of the IV estimator may be affected by whether we use the Rogaland dummy variable as the only instrument, or use multiple instruments by interacting the Rogaland dummy variable with some control variables. Given our large samples and strong first stage results reported in Table B1, small sample bias in the IV estimator because of multiple instruments is likely of little concern (see e.g., Staiger and Stock 1997). It is nevertheless reassuring to find in Table B3 that the quadratic IV results when using only two interaction instruments yields very similar results as those reported in Table 4. Moreover, we see that the quadratic IV estimates change little when interacting the Rogaland dummy variable with each included control variable.

Log Specification.—Above, we have followed Dahl and Lochner (2011) and others in measuring family income in levels rather than logs. However, some previous studies have preferred to specify family income as a linear function of log income, perhaps to allow for the effect of income to be stronger in poor families. For comparison and as a robustness check of our results, Table B4 shows results when replacing family income in levels with logs (row d) in Model 1 (excluding families with zero family income). In line with Dahl and Lochner (2011), we find that measuring family in logs rather than levels does not change our main results. Specifically, the linear OLS estimates indicate a positive and sizable effect of family income on children's outcomes. For example, a 10 percent increase in family income corresponds to an additional 0.1 years of education. Moreover, the linear FE estimates of family income are close to zero, and significantly different from the linear OLS estimates. We also see that the linear IV estimates show no sign of significant effects of family income on child outcomes, although they are too imprecisely estimated to rule out some effect. We conclude therefore that neither the linear specification in log family income nor the linear specification in level of family income is able to uncover the significant effect of family income on child outcomes.

Zero Income Families.—In an additional robustness check, we make sure that our results are not driven by the small number of observations with zero family income. From the linear estimates in Table B4, we see that results excluding families with zero family income (row b) give very similar results to the baseline estimates (row a), reported in Table 3. We do the same comparison for the quadratic estimates in Table B5. Also in this case, we find that the results without observations with zero income (row b) are very similar to the baseline estimates (row a) reported in Table 4.

Definition of Family Income.—We followed Løken (2010) in using family income when the child was aged 1–11. Her reason was that the oil boom mostly affected income during this period, and extending the measure of family income beyond age 11 did not change the results. To make sure that this holds true also for our analysis, Table B4 reports linear estimates when family income is measured as the average annual family income from age 1 until age 16, whereas Table B5 makes the same robustness check for the quadratic specification. We see that the estimates barely move depending on how we measure family income.

IV. Conclusion

Empirical studies often use IV or FE methods to estimate causal relationships when concerns about endogeneity bias arise. Yet in many empirical applications, there is no particular reason to expect that the true relationship is linear. In this paper, we have examined the implications of nonlinearity for IV and FE estimation of the causal link between childhood family income and child outcomes. Our nonlinear IV and FE estimates show an increasing, concave relationship between family income and children's outcomes, as predicted by economic theory. We find that the linear estimators miss the significant effects of family income because they assign little weight to the large marginal effects in the lower part of the income distribution. We also show that the linear IV and FE estimates differ primarily because of different weighting of marginal effects.

A general lesson to be drawn from our study is that the linearity restriction in IV and FE estimation may drive the conclusions reached in applications where there are reasons to suspect a nonlinear relationship. In fact, IV and FE estimation may exacerbate the sensitivity of the results to functional form assumptions because of the way they weight the underlying marginal effects. The ideal remedy is sensitivity analysis, showing how the results vary with changes in functional form. And further, it is useful to compute the weight functions, to know what the linear estimates actually identify and interpret the results in view of that.

Our results may also be of interest from a policy perspective. Most developed countries have a range of policies targeted at family income during childhood, such as family allowances, maternity benefits, single parent benefits, and family tax credits. In fact, families with children receive special treatment under the tax-benefit system in 28 of the 30 OECD-countries (OECD 2002).[18] While some of

[18] See Del Boca, Flinn, and Wiswall (2010) for an analysis of the differential effects of these types of policies within an estimated household model of child development.

these policies are means-tested, others are more universal in nature. Our IV and FE results suggesting a concave relationship between family income and child outcomes, with relatively large positive effects in the lower part of the family income distribution, indicate that policies targeting poor families may be quite effective in promoting child development.

Appendix A

A. *FE as IV Estimator*

This Appendix shows that the FE estimator is numerically equal to an IV estimator using an appropriately chosen instrument.

The FE estimator is given by

$$\hat{\beta}(FE) = \frac{\sum_{j=1}^{J} \Delta C_j \Delta y_j - \frac{1}{J}\left(\sum_{j=1}^{J} \Delta C_j\right)\left(\sum_{j=1}^{J} \Delta y_j\right)}{\sum_{j=1}^{J} \Delta C_j \Delta C_j - \frac{1}{J}\left(\sum_{j=1}^{J} \Delta C_j\right)\left(\sum_{j=1}^{J} \Delta C_j\right)}.$$

Define the FE instrument in the pooled sample of siblings as

$$q_{jb} = 1\{b = 2\}\left(\Delta C_j - \frac{1}{J}\sum_{j=1}^{J} \Delta C_j\right) - 1\{b = 1\}\left(\Delta C_j - \frac{1}{J}\sum_{j=1}^{J} \Delta C_j\right).$$

For a sample of J families and $2J$ observations, we can define the IV estimator $\hat{\beta}(q)$ using the q_{jb} instrument as

$$\hat{\beta}(q) = \frac{\sum_{b=1}^{2} \sum_{j=1}^{J} y_{jb} q_{jb}}{\sum_{b=1}^{2} \sum_{j=1}^{J} C_{jb} q_{jb}},$$

since $\sum_{b=1}^{2} \sum_{j=1}^{J} q_{jb} = 0$.

Substituting the instrument and rearranging, we have

$$\hat{\beta}(q) = \frac{\sum_{j=1}^{J} \{- y_{j1} \Delta C_j + y_{j2} \Delta C_j + y_{j1}\frac{1}{J}\sum_{j=1}^{J} \Delta C_j - y_{j2}\frac{1}{J}\sum_{j=1}^{J} \Delta C_j\}}{\sum_{j=1}^{J} \{- C_{j1} \Delta C_j + C_{j2} \Delta C_j + C_{j1}\frac{1}{J}\sum_{j=1}^{J} \Delta C_j - C_{j2}\frac{1}{J}\sum_{j=1}^{J} \Delta C_j\}}$$

$$= \frac{\sum_{j=1}^{J} (y_{j2} - y_{j1})\Delta C_j - \frac{1}{J}\left(\sum_{j=1}^{J} (y_{j2} - y_{j1})\right)\left(\sum_{j=1}^{J} \Delta C_j\right)}{\sum_{j=1}^{J} (C_{j2} - C_{j1})\Delta C_j - \frac{1}{J}\left(\sum_{j=1}^{J} (C_{j2} - C_{j1})\right)\left(\sum_{j=1}^{J} \Delta C_j\right)}$$

$$= \hat{\beta}(FE).$$

B. *Decomposition of Linear Estimators with Heterogeneous Marginal Effects*

With heterogeneous marginal effects, we can generalize the unrestricted model (2) as follows:

$$\text{(A1)} \qquad y_i = \mu + \sum_{c=1}^{c} \gamma_{ci} d_{ci} + v_i,$$

where γ_{ci} is a random coefficient, capturing that two individuals may have different effects of a one unit increase from family income level $c - 1$ to c, that is $\gamma_{ci} \neq \gamma_{ci'}$ for $i \neq i'$. This model nests the unrestricted model, which assumes homogenous marginal effects: $\gamma_{ci} = \gamma_c$ for all i. The linear model (1) assumes that the marginal effects are homogenous and independent of family income level: $\gamma_{ci} = \beta$ for all c and i.

With heterogenous marginal effects, the linear OLS, IV, and FE estimators will identify different weighted averages of different marginal effects estimates. While the expressions for the weights remain as given in (3), (4), and (8), the expressions for the marginal effects estimands need to be modified.

The OLS estimand for β in (1) is $\beta(OLS) = \text{Cov}(y_i, C_i)/\text{Var}(C_i)$. By substituting for y_i, we can write the linear OLS estimand as

$$\text{(A2)} \qquad \beta(OLS) = \sum_{c=1}^{c} \gamma_c(OLS) w_c(OLS),$$

where the OLS estimand for the cth marginal effect is

$$\gamma_c(OLS) = E[\gamma_{ci} | C_i = c] - E[\gamma_{ci} | C_i = c - 1] + \Delta_c,$$

and $\Delta_c = E[v_i | C_i = c] - E[v_i | C_i = c - 1]$ represents the selection bias in $\gamma_c(OLS)$. The OLS estimand for the cth marginal effect is then the difference in the average outcome for individuals with $C_i = c$ and $C_i = c - 1$. The associated OLS weight remains as given in (3).

The IV estimand for β in (1) is $\beta(Z) = \text{Cov}(y_i, Z_i)/\text{Cov}(C_i, Z_i)$. By substituting for y_i, we get

$$\beta(Z) = \sum_{c=1}^{c} \frac{\text{Cov}(d_{ci}\gamma_{ci}, Z_i)}{\text{Cov}(C_i, Z_i)}.$$

Without loss of generality, let $E[Z_i] = 0$, and rewrite the expression as

$$\beta(Z) = \sum_{c=1}^{\bar{c}} \frac{E[\gamma_{ci} Z_i | d_{ci} = 1] \operatorname{pr}(d_{ci} = 1)}{\text{Cov}(C_i, Z_i)}.$$

Then, substituting the weight expression from (4), we have

$$\beta(Z) = \sum_{c=1}^{\bar{c}} w_c(Z)\gamma_c(Z), \tag{A3}$$

where the IV estimand for the cth marginal effect is given by

$$\gamma_c(Z) = \frac{E[\gamma_{ci}Z_i \,|\, d_{ci} = 1]}{E[Z_i \,|\, d_{ci} = 1]}.$$

Note that unlike the homogenous case, the IV estimand for the cth marginal effect depends on the particular instrument. We therefore label the marginal effect $\gamma_c(Z)$ rather than γ_c. With homogenous marginal effects, then

$$\gamma_c(Z) = \frac{E[\gamma_c Z_i \,|\, d_{ci} = 1]}{E[Z_i \,|\, d_{ci} = 1]} = \gamma_c.$$

This means that the linear IV estimand with heterogenous marginal effects can be interpreted as a weighted average of the marginal local average treatment effects (see Angrist and Imbens 1995).

In the panel data setting, we can generalize the model to allow for heterogeneity by indexing the marginal effect by family j and birth order b:

$$y_{jb} = f_{jb}(C_{jb}) = \mu_{jb} + \sum_{c=1}^{\bar{c}} \gamma_{cjb} d_{cjb} + \upsilon_{jb}. \tag{A4}$$

The linear FE estimand for β in (6) is $\beta(FE) = \mathrm{Cov}(\Delta y_j, \Delta C_j)/\mathrm{Var}(\Delta C_j)$. Given that the FE estimator is numerically equal to an IV estimator using the q_{jb} instrument, we get

$$\beta(FE) = \sum_{c=1}^{\bar{c}} w_c(FE)\gamma_c(FE), \tag{A5}$$

where

$$\gamma_c(FE) = \frac{E[\gamma_{cjb} q_{jb} \,|\, d_{cjb} = 1]}{E[q_{jb} \,|\, d_{cjb} = 1]}.$$

Note that unlike the homogenous case, the FE estimand for the cth marginal effect will depend on the subsample of sibling pairs that experience within-family income variation at this family income level. We therefore label the marginal effect $\gamma_c(FE)$ rather than γ_c.

As is evident from equations (A2), (A3), and (A5), with heterogeneity in marginal effects the linear OLS, IV, and FE estimators will identify different weighted averages of different marginal effects estimates. The marginal effects estimates differ in part because they represent the average marginal effects for different subgroups, but also because of endogeneity bias in the OLS estimation.

Appendix B: Tables

Table B1—First Stage Regression Results

	Instruments							
	Born in Rogaland in 1967–1969	Predicted income	Predicted income squared	Rog* mother college	Rog* father college	Rog* mother age	Rog* father age	Rog* large family
Panel A. Linear IV								
Family income in NOK 10,000	0.757***							
	(0.170)							
Panel B. Quadratic IV								
Family income in NOK 10,000		−0.418***	0.030***					
		(0.128)	(0.002)					
Squared family income		−55***	2.19***					
		(7.9)	(0.117)					
Panel C.								
Family income in NOK 10,000	1.02***			−2.50***	−1.59***	−0.091	0.340	0.998***
	(0.186)			(0.448)	(0.345)	(0.211)	(0.286)	(0.269)
Squared family income	75***			−138***	−71***	−0.51	2.6	30.0**
	(9.8)			(29)	(14)	(19)	(12)	(15)

Notes: This table reports first stage results from the IV estimation of Model 1 and 2. Panel A uses the born in Rogaland in 1967–1969 dummy variable instrument. Panel B uses the set of interacted instruments (born in Rogaland in 1967–1969 dummy variable and interacting the born in Rogaland in 1967–1969 dummy variable with father's college, mother's college, father's age, mother's age and large family) in First Stages I and II, whereas panel C uses the predicted family income instruments based on the same set of interacted instruments in First Stages I and II. A full set of controls is used in all regressions. Note that the quadratic income regressor is linear income (in NOK 10,000) multiplied by 100, hence quadratic income is in NOK 1,000,000. Standard errors in parentheses are heteroskedastic robust. F-test on excluded instruments is 190 (panel B, Family Income in NOK 10,000), 211 (panel B, Squared family income, 21.6 (panel C, Family income in NOK 10,000), and 19.8 (panel C, Squared family income). The F-test is performed on the excluded instruments (and not the controls).

***Significant at the 1 percent level. **Significant at the 5 percent level *Significant at the 10 percent level.

Table B2—Linear IV Estimates with Different Instruments

	Dependent variables		
	Education	Dropout	Adult IQ (males only)
Panel A. Linear IV			
Instrument: born in Rogaland in 1967–1969			
Family income in NOK 10,000	0.022	−0.012	0.033
	(0.057)	(0.011)	(0.023)
Panel B. Linear IV			
Instruments: interactions			
Family income in NOK 10,000	0.026	−0.008	−0.017
	(0.020)	(0.006)	(0.013)
Panel C. Linear IV			
Instruments: predicted income	−0.010	0.001	−0.018
Family income in NOK 10,000	(0.012)	(0.001)	(0.013)
Observations	121,122	121,122	57,788

Notes: This table reports IV estimates of Model 1 using the IV sample. Panel A uses the born in Rogaland in 1967–1969 dummy variable instrument. Panel B uses the set of interacted instruments (born in Rogaland in 1967–1969 dummy variable and interacting the born in Rogaland in 1967–1969 dummy variable with father's college, mother's college, father's age, mother's age and large family), whereas panel C uses the predicted family income instruments based on the same set of interacted instruments. A full set of controls is used in all regressions. Standard errors in parentheses are heteroskedastic robust.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

TABLE B3—QUADRATIC IV WITH DIFFERENT INTERACTION INSTRUMENTS

	Dependent variables		
	Education	Dropout	Adult IQ (males only)
Panel A. Quadratic IV			
Instruments: all interactions			
Family income in NOK 10,000	0.178**	−0.046***	0.171**
	(0.075)	(0.014)	(0.079)
Quadratic income (× 100)	−0.260*	0.067**	−0.246*
	(0.139)	(0.022)	(0.149)
Panel B. Quadratic IV			
Instruments: predicted income			
Family income in NOK 10,000	0.199***	−0.070***	0.241***
	(0.071)	(0.013)	(0.070)
Quadratic income (× 100)	−0.301***	0.113***	−0.382***
	(0.106)	(0.019)	(0.107)
Panel C. Quadratic IV			
Instruments: interactions			
Family income in NOK 10,000	0.207**	−0.036*	0.256**
	(0.100)	(0.019)	(0.114)
Quadratic income (× 100)	−0.316*	0.045	−0.432*
	(0.180)	(0.035)	(0.222)
Panel D. Quadratic IV			
Instruments: predicted income			
Family income in NOK 10,000	0.192**	−0.072***	0.212***
	(0.088)	(0.015)	(0.071)
Quadratic income (× 100)	−0.288**	0.114***	−0.342***
	(0.124)	(0.022)	(0.105)
Observations	121,122	121,122	57,788

Notes: This table reports IV estimates of Model 2. Panel A uses a full set of interacted instruments in First stages I and II (born in Rogaland in 1967–1969 dummy variable and interacting the born in Rogaland in 1967–1969 dummy variable with all included covariates). Panel B uses the predicted family income instruments based on this full set of interacted instruments in First Stages I and II. Panel C uses a subset of interacted instruments (born in Rogaland in 1967–1969 dummy variable and interacting the born in Rogaland in 1967–1969 dummy variable with only mother's college and large family) in First Stages I and II. Panel D uses the predicted family income instruments based on this subset of interacted instruments in First Stages I and II. A full set of controls is used in all regressions. Note that the quadratic income regressor is linear income (in NOK 10,000) multiplied by 100, hence quadratic income is in NOK 1,000,000. Standard errors in parentheses are heteroskedastic robust.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

TABLE B4—LINEAR OLS, IV, AND FE ESTIMATES—ROBUSTNESS TESTS II

	Education	Dropout	Adult IQ (males only)
Panel A. Linear OLS			
a) Baseline	0.043***	−0.006***	0.028***
	(0.001)	(0.001)	(0.001)
b) Drop 0 income	0.043***	−0.006***	0.029***
	(0.001)	(0.000)	(0.001)
c) Income 1–16	0.043***	−0.006***	0.028***
	(0.001)	(0.000)	(0.001)
d) ln(income)	0.010***	−0.001***	0.006***
	(0.000)	(0.000)	(0.000)
Panel B. Linear IV			
a) Baseline	0.022	−0.012	−0.061
	(0.056)	(0.011)	(0.072)
b) Drop 0 income	0.018	−0.013	−0.062
	(0.061)	(0.012)	(0.073)
c) Income 1–16	0.023	−0.014	−0.076
	(0.069)	(0.014)	(0.091)
d) ln(income)	0.007	−0.006	−0.024
	(0.027)	(0.006)	(0.027)
Panel C. Linear OLS			
a) Baseline	0.041***	−0.005***	0.024***
	(0.001)	(0.000)	(0.001)
b) Drop 0 income	0.041***	−0.005***	0.023***
	(0.001)	(0.000)	(0.001)
c) Income 1–16	0.043***	−0.006***	0.025***
	(0.001)	(0.000)	(0.001)
d) ln(income)	0.010***	−0.001***	0.006***
	(0.000)	(0.000)	(0.000)
Panel D. Linear FE			
a) Baseline	0.000	−0.001	−0.001
	(0.003)	(0.001)	(0.004)
b) Drop 0 income	0.000	−0.001	−0.001
	(0.004)	(0.001)	(0.004)
c) Income 1–16	0.003	−0.001	0.001
	(0.003)	(0.001)	(0.004)
d) ln(income)	0.001	−0.000	0.001
	(0.001)	(0.000)	(0.001)

Notes: This table reports OLS, IV, and FE estimates of Model 1. Panels A and B use the IV sample, whereas panel C and D uses the FE sample. Panel B uses first born in Rogaland in 1967–1969 dummy variable as the only instrument. A full set of controls is used in all regressions. Standard errors in parentheses are heteroskedastic robust.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table B5—Quadratic OLS, IV, and FE—Robustness Tests III

	Education		Dropout		Adult IQ (males only)	
Panel A. Quadratic OLS						
a) Baseline	0.051***	(0.002)	−0.010***	(0.000)	0.032***	(0.002)
	−0.016***	(0.005)	0.008***	(0.001)	−0.008*	(0.004)
b) Drop 0 income	0.052***	(0.002)	−0.010***	(0.001)	0.033***	(0.003)
	−0.018***	(0.002)	0.008***	(0.001)	−0.010***	(0.005)
c) Income 1–16	0.050***	(0.002)	−0.010***	(0.000)	0.029***	(0.002)
	−0.012***	(0.004)	0.008***	(0.001)	−0.002	(0.004)
Panel B. Quadratic IV						
Instruments: Interactions						
a) Baseline	0.180**	(0.087)	−0.030*	(0.016)	0.234**	(0.109)
	−0.302*	(0.164)	0.042*	(0.021)	−0.401**	(0.210)
b) Drop 0 income	0.174**	(0.085)	−0.028*	(0.015)	0.211*	(0.109)
	−0.293*	(0.160)	0.039	(0.030)	−0.363*	(0.213)
c) Income 1–16	0.162**	(0.077)	−0.027**	(0.014)	0.203**	(0.086)
	−0.266*	(0.145)	0.036	(0.028)	−0.357**	(0.170)
Panel C. Quadratic IV						
Instruments: Predicted income						
a) Baseline	0.142**	(0.072)	−0.057***	(0.013)	0.195***	(0.070)
	−0.228**	(0.107)	0.097***	(0.019)	−0.323***	(0.106)
b) Drop 0 income	0.144**	(0.068)	−0.054***	(0.011)	0.168***	(0.063)
	−0.235**	(0.097)	0.091***	(0.016)	−0.283***	(0.090)
c) Income 1–16	0.136*	(0.084)	−0.059***	(0.015)	0.166**	(0.076)
	−0.213*	(0.127)	0.101***	(0.023)	−0.278**	(0.113)
Panel D. Quadratic OLS						
a) Baseline	0.069***	(0.002)	−0.012***	(0.000)	0.033***	(0.003)
	−0.050***	(0.003)	0.013***	(0.001)	−0.017***	(0.004)
b) Drop 0 income	0.069***	(0.002)	−0.013***	(0.000)	0.033***	(0.003)
	−0.050***	(0.003)	0.013***	(0.001)	−0.017***	(0.004)
c) Income 1–16	0.071***	(0.002)	−0.014***	(0.000)	0.033***	(0.003)
	−0.047***	(0.003)	0.013***	(0.001)	−0.013***	(0.004)
Panel E. Quadratic FE						
a) Baseline	0.041***	(0.008)	−0.006***	(0.002)	0.019*	(0.010)
	−0.065***	(0.013)	0.008***	(0.002)	−0.031*	(0.014)
b) Drop 0 income	0.042***	(0.008)	−0.006***	(0.002)	0.013	(0.012)
	−0.065***	(0.013)	0.008***	(0.002)	−0.018	(0.016)
c) Income 1–16	0.055***	(0.007)	−0.009***	(0.001)	0.022**	(0.009)
	−0.080***	(0.010)	0.012***	(0.002)	−0.033***	(0.012)

Notes: This table reports OLS, IV, and FE estimates of Model 2. Panels A and B uses the IV sample, whereas panels C and D use the FE sample. The top row in each specification is the linear coefficient on family income (in NOK 10,000), and the bottom row is the quadratic coefficient ($\times$ 1,000). Panel B uses the set of interacted instruments (born in Rogaland in 1967–1969 dummy variable and interacting the born in Rogaland in 1967–1969 dummy variable with father's college, mother's college, father's age, mother's age, and large family) in First Stages I and II, whereas panel C uses the predicted family income instruments based on the same set of interacted instruments in First Stages I and II. A full set of controls is used in all regressions. Note that the quadratic income is linear income (in NOK 10,000) multiplied by 100, hence quadratic income is in NOK 1,000,000. Standard errors in parentheses are heteroskedastic robust.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

REFERENCES

Almond, Douglas, and Janet Currie. 2010. "Human Capital Development Before Age Five." National Bureau of Economic Research Working Paper 15827.

Angrist, Joshua D., Kathryn Graddy, and Guido W. Imbens. 2000. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish." *Review of Economic Studies* 67 (3): 499–527.

Angrist, Joshua D., and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430): 431–42.

Angrist, Joshua D., and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*. Vol. 3A, edited by Orley Ashenfelter and David Card, 1277–1366. Amsterdam: Elsevier Science.

Becker, Gary S., and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87 (6): 1153–89.

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4 (3): S1–39.

Blau, David M. 1999. "The Effect of Income on Child Development." *Review of Economics and Statistics* 81 (2): 261–76.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101(6): 27–81.

Dahl, Gordon B., and Lance Lochner. 2011. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." dss.ucsd.edu/~gdahl/papers/children-and-EITC.pdf.

Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2010. "Household Choices and Child Development." Institute for the Study of Labor (IZA) Discussion Paper 5155.

Dooley, Martin, and Jennifer Stewart. 2004. "Family Income and Child Outcomes in Canada." *Canadian Journal of Economics* 37 (4): 898–917.

Duncan, Greg J., W. Jean Yeung, Jeanne Brooks-Gunn, and Judith R. Smith. 1998. "How Much Does Childhood Poverty Affect the Life Chances of Children?" *American Sociological Review* 63 (3): 406–23.

Eide, Martha G., Nina Øyen, Rolv Skjærven, Stein Tore Nilsen, Tor Bjerkedal, and Grethe S. Tell. 2005. "Size at Birth and Gestational Age as Predictors of Adult Height and Weight." *Epidemiology* 16 (2): 175–81.

Heckman, James J. 2008. "Econometric Causality." *International Statistical Review* 76 (1): 1–27.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2008. "Earnings Functions and Rates of Return." *Journal of Human Capital* 2 (1): 1–31.

Heckman, James J., and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73 (3): 669–738.

Horowitz, Joel L. 2011. "Applied Nonparametric Instrumental Variables Estimation." *Econometrica* 79 (2): 347–94.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.

Levy, Dan, and Greg Duncan. 2000. "Using Sibling Samples to Assess the Effect of Childhood Family Income on Completed Schooling." Northwestern University, Joint Center for Policy Research (JCPR) Working Paper 168.

Lochner, Lance, and Enrico Moretti. 2001. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." National Bureau of Economic Research Working Paper 8605.

Lochner, Lance, and Enrico Moretti. 2011. "Estimating and Testing Non-Linear Models Using Instrumental Variables." National Bureau of Economic Research Working Paper 17039.

Løken, Katrine V. 2010. "Family Income and Children's Education: Using the Norwegian Oil Boom as a Natural Experiment." *Labour Economics* 17 (1): 118–29.

Løken, Katrine V., Magne Mogstad, and Matthew Wiswall. 2012. "What Linear Estimators Miss: The Effects of Family Income on Child Outcomes: Dataset." *American Economic Journal: Applied Economics*. http://dx.doi.org/10.1257/app.4.2.1.

Mayer, Susan E. 1997. *What Money Can't Buy: Family Income and Children's Life Chances*. Cambridge, MA: Harvard University Press.

Milligan, Kevin, and Mark Stabile. 2007. "The Integration of Child Tax Credits and Welfare: Evidence from the Canadian National Child Benefit Program." *Journal of Public Economics* 91 (1–2): 305–26.

Mogstad, Magne, and Matthew Wiswall. 2011a. "Testing the Quantity-Quality Model of Fertility: Estimation Using Unrestricted Family Size Models." http://files.nyu.edu/mw109/public/testing_qq.pdf

Mogstad, Magne, and Matthew Wiswall. 2011b. "Linearity in Instrumental Variables Estimation: Problems and Solutions." http://files.nyu.edu/mw109/public/mogstad_wiswall_IV_linearity.pdf

Newey, Whitney K., and James L. Powell. 2003. "Instrumental Variable Estimation of Nonparametric Models." *Econometrica* 71 (5): 1565–78.

Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens. 2008. "The Intergenerational Effects of Worker Displacement." *Journal of Labor Economics* 26 (3): 455–83.

Organisation for Economic Co-Operation and Development (OECD). 2002. *Taxing Wages, 2001 Edition*. Washington, DC: OECD.

Robinson, Peter M. 1988. "Root-N-Consistent Semiparametric Regression." *Econometrica* 56 (4): 931–54.

Shea, John. 2000. "Does Parents' Money Matter?" *Journal of Public Economics* 77 (2): 155–84.

Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*. Vol. 3A, edited by Orley Ashenfelter and David Card, 1761–1800. Amsterdam: Elsevier Science.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65 (3): 557–86.

Sundet, John Martin, Dag G. Barlaug, and Tore M. Torjussen. 2004. "The End of the Flynn Effect?: A Study of Secular Trends in Mean Intelligence Test Scores of Norwegian Conscripts During Half a Century." *Intelligence* 32 (4): 349–62.

Sundet, John Martin, Kristian Tambs, Jennifer R. Harris, Per Magnus, and Tore M. Torjussen. 2005. "Resolving the Genetic and Environmental Sources of the Correlation Between Height and Intelligence: A Study of Nearly 2600 Norwegian Male Twin Pairs." *Twin Research and Human Genetics* 8 (4): 307–11.

Thrane, Vidkunn Coucheron. 1977. "Evneprøving av Utkrivingspliktige i Norge 1950–1953." Arbeidsrapport nr. 26, INAS.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* 2nd ed. Cambridge, MA: MIT Press.