DISCUSSION PAPER SERIES

IZA DP No. 9605

Lower Bounds and the Linearity Assumption in Parametric Estimations of Inequality of Opportunity

Paul Hufe
Andreas Peichl

December 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Lower Bounds and the Linearity Assumption in Parametric Estimations of Inequality of Opportunity

Paul Hufe
ZEW Mannheim
and University of Mannheim

Andreas Peichl
ZEW Mannheim,
University of Mannheim, IZA and CESifo

Discussion Paper No. 9605
December 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9605
December 2015

ABSTRACT

Lower Bounds and the Linearity Assumption in Parametric Estimations of Inequality of Opportunity*

The consistent underestimation of inequality of opportunity has led some scholars to call into question the usefulness of such estimates. In this paper we argue that neglecting heterogeneity in the influence of circumstances across types as well as neglecting heterogeneity in type-specific effort distributions are two important sources of the downward bias in inequality of opportunity measures. Compared to the standard parametric approach of ex ante measurement of inequality of opportunity, we calculate a 50% upwards correction when accounting for both sources of heterogeneity. Therefore, taking heterogeneity across types seriously is an important step towards strengthening the policy relevance of this concept.

JEL Classification: D63, D3

Keywords: equality of opportunity, type heterogeneity, parametric estimation

Corresponding author:

Andreas Peichl
Centre for European Economic Research (ZEW)
L7,1
68161 Mannheim
Germany
E-mail: peichl@zew.de

* We are grateful to John Roemer and Martin Ungerer for valuable comments and suggestions.

1 Introduction

There is now a large theoretical and empirical literature in economics on inequality of opportunity (IOp).[1] In one prominent formulation (Roemer, 1993, 1998), outcomes that individuals enjoy (such as income) are consequences of two sorts of factor: Circumstances, those characteristics of a person and her environment that are beyond her control or for which she should not be held responsible, and effort, which comprises those choices within her realm of control. Equality of opportunity is said to hold when the chances that individuals face for achieving the outcome in question are independent of their circumstances, and sensitive only to personal effort.

Following the work of Bourguignon et al. (2007) and Ferreira and Gignoux (2011) many scholars have favored parametric estimations of IOp over the non-parametric approach, e.g. as used in Checchi and Peragine (2010). Somewhat surprisingly, however, applied works using parametric estimations are reluctant to incorporate type dependent heterogeneity in the influence of particular circumstances and efforts. First, instead of including interaction terms, circumstance variables are introduced linearly, which by necessity implies that researchers assume a homogeneous influence of circumstances across the partition of Roemerian types. Second, it is recognized that the distribution of efforts is itself type-dependent. Therefore, leaving residuals from parametric IOp estimations unstandardized must be based on the presumption that the obtained distribution of type-specific error terms is indicative for an ethically non-objectionable effort distribution (Roemer and Trannoy, 2015). It is well understood that IOp estimates are downward biased in case of unobserved circumstances (Balcázar, 2015; Ferreira and Gignoux, 2011; Niehues and Peichl, 2014), perhaps importantly so. In addition to constraints in data availability, neglecting both sources of type-specific heterogeneity may be important sources for that bias.

In what follows we outline how neglecting type-specific heterogeneity re-enforces the underestimation of IOp in the context of imperfect information on the relevant set of circumstances. Section 3 demonstrates the magnitude of this underestimation using data from the Child & Young Adults Supplement of the National Longitudinal Survey of Youth (NLSY79). Section 4 concludes.

[1] For recent surveys, see Ramos and Van de gaer (2016); Roemer and Trannoy (2015), or Ferreira and Peragine (2015).

2 Parametric IOp Estimation

The literature on IOp commonly assumes that a set of circumstances Ω and a scalar θ of effort determine the outcome of interest p. The relation between these components can be described by a function $g : \theta \times \Omega \mapsto \mathbb{R}^+$. As it appears reasonable to assume that the distribution of efforts is not orthogonal to circumstances the relation of interest can be rewritten in the following form:

$$p = g(\Omega, \theta(\Omega), \epsilon) \tag{1}$$

where circumstances Ω are considered as root causes of unfair inequality beyond individual control, whereas differential effort θ net of circumstance influence yields fair inequality. Based on the realizations x_j of each circumstance $C^j \in \Omega$ we can partition the population into a set of types T, where the number of types is given by $K = \prod_{j=1}^{J} x_j$. According to the ex ante approach of measuring IOp, perfect equality of opportunity would prevail if the type specific mean advantage levels $\mu^k(p)$ were equal across all types $T^k \in T$. Thus, the degree of inequality in a smoothed distribution Φ, in which each individual income p_i is replaced by the mean income of the respective type $\mu^k(p)$ can be considered as a measure of IOp. The share of unfair inequalities in the aggregate distribution of advantages $F(p)$ would be given by

$$IOR = I(\Phi)/I(F(p)). \tag{2}$$

According to the standard parametric approach the distribution of $\mu^k(p)$ would now be constructed in two steps:

$$\ln p_i = \beta_0 + \sum_{j=1}^{J} \beta_j C_i^j + \epsilon_i \tag{3}$$

$$\mu^k(p) = \exp\left[\sum_{j=1}^{J} \hat{\beta}_j C_i^j\right] \tag{4}$$

By definition $C_i = C_j$, $\forall i, j \in T^k$ and thus the predicted values from (4) yield K type-specific averages. Yet it is noteworthy that the coefficients β_j are independent of type T^k. Thus, any heterogeneity in β_j is implicitly attributed to the residual.

The non-parametric approach advanced by Checchi and Peragine (2010) would simply average advantage levels within types. Note that the same operation can be executed within the parametric framework outlined above by regressing the outcome of interest on the intercept and a set of $K-1$ group dummies:

$$\ln p_i = \beta_0 + \sum_{k=1}^{K-1} \beta_k \mathbb{1}(i \in T^k) + \epsilon_i \tag{5}$$

$$\mu^k(p) = \exp\left[\sum_{k=1}^{K-1} \hat{\beta}_k \mathbb{1}(i \in T^k)\right] \tag{6}$$

The first approach described in equations (3) and (4) would only yield an unbiased estimate $\mu^k(p)$ if the effect of each C^j was indeed homogeneous across all types T^k. To illustrate this fact in an intuitive manner, consider the simple case of two binary circumstances, say sex and the high-school graduation status of the respondent's mothers, which in turn yields the following type partition T:

	Male	Female
Non-Graduate Mother	Type 1	Type 2
Graduate Mother	Type 3	Type 4

The non-linear case would be estimated as follows:

$$\ln p_i = \beta_1 + \beta_2 C_i^{female} + \beta_3 C_i^{HS} + \beta_4 [C_i^{female} \times C_i^{HS}] + \epsilon_i \tag{7}$$

Note that equation (7) is equivalent to (5) as each parameter β_k can be interpreted as the natural logarithm of $\mu^k(p)$, i.e. the type-specific mean advantage level. The standard approach in the literature, corresponding to equation (3), however, reads as follows:

$$\ln p_i = \beta_1 + \beta_2 C_i^{female} + \beta_3 C_i^{HS} + \tilde{\epsilon}_i \tag{8}$$

$$= \beta_1 + \beta_2 C_i^{female} + \beta_3 C_i^{HS} + (\epsilon_i + \beta_4 [C_i^{female} \times C_i^{HS}]) \tag{9}$$

Clearly, the two approaches do only coincide in case of $\beta_4 = 0$; in our example if the influence of maternal education was homogeneous across gender types (or vice versa). Only then, the additive-linear introduction of circumstances would be warranted.

To put it in general terms, the standard approach (Equations (3), (4)) would deliver unbiased estimates of $\mu^k(p)$ if the circumstance influence was driven by the coefficients of the non-interacted base levels C^j, only. To the contrary, if the coefficients on the interaction terms were non-zero we would underestimate IOp by attributing type-specific heterogeneity in coefficients to the error term.

Furthermore, it is reasonable to assume that the influence of effort is heterogeneous across types. The difference is partially taken into account by estimating the influence of circumstances on outcomes in a reduced form akin to equation (1). Note, however, that the reduced-form only nets out type-specific heterogeneity in effort *levels*. What is not captured is differences in *within-type effort variance*. According to the Roemerian approach to IOp, one may argue that it is beyond individual control in which circumstance group individuals are born. Therefore, one shall not be held accountable for the type-dependent set of potential efforts. Conditional on accepting the underlying normative assumptions, it follows that heteroskedasticity across circumstance must be modeled explicitly to yield standardized residual (and therefore effort) distributions. A procedure to standardize residuals has been suggested by Björklund et al. (2012), who found that the type-dependent variance in effort levels provided a substantial source of IOp. To be precise, they calculated standardized effort levels, u_i, with type-independent variance as follows: $u_i = \epsilon_i \sigma / \sigma_{T^k}$, where σ and σ_{T^k} indicate the overall standard deviation of error terms and the type-specific error standard deviation, respectively. In order to keep group-specific mean outcomes unaffected, equation (6) has to be re-written as follows:

$$\mu^k(p) = \exp\left[\sum_{k=1}^{K} \hat{\beta}_k \mathbb{1}(i \in T^k) + \epsilon_i - \underbrace{\epsilon_i \sigma / \sigma_{T^k}}_{=u_i}\right] \tag{10}$$

From equation (10) it follows intuitively, that (6) would yield biased estimates of IOp if $\exists\ \sigma_{T^k} \neq \sigma$. Only if all type-specific error distributions were homoskedastic both approaches would coincide. In terms of the implementation, Björklund et al. (2012) suggest to regress type-specific variances on the set of circumstance variables and to calculate u_i based on predicted values in order to smooth out the strong influence from types with extremely small variances. We adhere to their advice in what follows. To investigate the empirical relevance of neglecting both sources of heterogeneity we will now turn to the empirical application.

3 Application

For the sake of this illustration, we use the Child & Young Adults Supplement of the National Longitudinal Survey of Youth (NLSY79). The outcome of interest p is gross income averaged over the age range 25 to 30 (see Table 1 for summary statistics). We consider five circumstance variables, which are sequentially introduced to yield five circumstance scenarios. First, we consider the respondent's sex, with female being the omitted category

Table 1: Summary Statistics

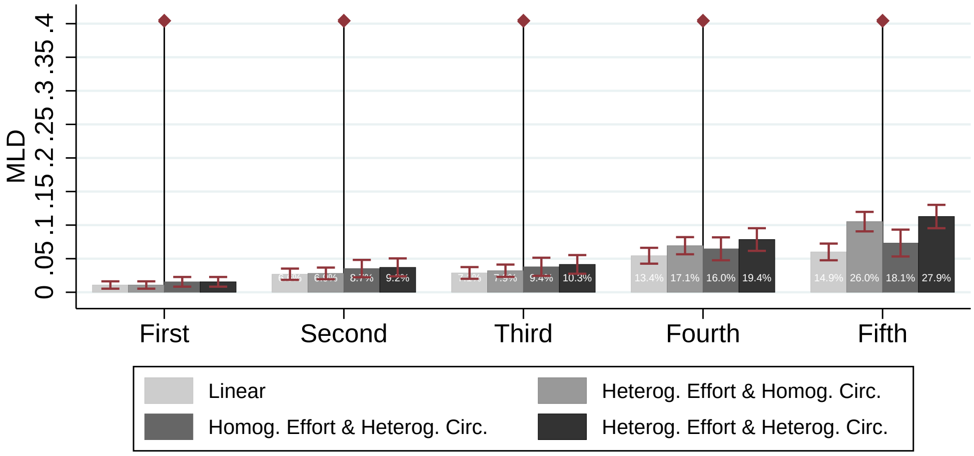
	(1) N	(2) mean	(3) sd	(4) min	(5) max
Avrg. Prim. Income (25-30)	3,149	24,871	20,872	22.50	247,655
Male	3,149	0.493	0.500	0	1
Majority	3,149	0.430	0.495	0	1
Avrg. Net Fam. Inc. <P25	3,149	0.277	0.448	0	1
Avrg. Net Fam. Inc. <P50	3,149	0.287	0.453	0	1
Avrg. Net Fam. Inc. <P75	3,149	0.277	0.447	0	1
Secondary (16)	3,149	0.616	0.486	0	1
Intermediate (16)	3,149	0.0959	0.295	0	1
College (16)	3,149	0.109	0.312	0	1
SMSA, Not center (16)	3,149	0.508	0.500	0	1
SMSA, Center (16)	3,149	0.266	0.442	0	1

of the respective binary indicator. Second, we add the child's race by including a dummy variable, which takes on value one if the respondent is neither hispanic nor black. Third, we proxy for the respondent's residential environment at age 16 with a tripartite variable indicating whether the respondent lived in a Metropolitan Statistical Area (MSA), and if yes, whether she lived in the center of an MSA. The variable is dummified, with "living in a rural area" being the omitted category. Fourth, for all respondents we calculate the average household income of their family from birth until the age of 16. Families are then grouped into quartiles of this income distribution. We omit the highest income quartile from the resulting set of family income dummies. Lastly, we consider the academic achievement of respondent's mothers in four categories: high-school dropout, high-school graduate, intermediate post-secondary education and college graduate. We again create dummy variables and omit the high-school dropout category.

It follows that the sample can be partitioned in 192 non-overlapping types. Naturally, the NLSY79 allows for much larger circumstance sets. However, for the sake of this exposition we confine ourselves to a rather scant circumstance set in order to demonstrate the importance of non-linearities in contexts of poor data availability.

To investigate the impact of non-linearities on the estimated share of IOp in observed inequality, we contrast the results from the standard parametric approach to the results from the alternative non-linear approach. As previously mentioned the latter are derived by fully interacting the set of circumstances and thus account for type-specific heterogeneity in the influence of all C^j. Furthermore we adopt the standardization procedure suggested by (Björklund et al., 2012) to adjust our measure of IOp for heterogeneity in efforts. For the exposition of the results we will focus on the most extensive circumstance set unless we indicate otherwise.

Figure 1: Comparison linear vs. non-linear introduction of circumstances



Note: The central spike yields the extent of outcome inequality IO. The adjacent black colored bars of each each spike yield inequality attributed to circumstances, i.e. the lower bound absolute measure of inequality of opportunity IOp. On the leftmost side circumstances are introduced linearly, whereas on the center-left, the model is fully interacted. The center-right shows a model with linear circumstances but standardized efforts. The rightmost bar combines both sources of heterogeneity. The red whiskers indicate 95% confidence intervals, which are bootstrapped with 1000 repetitions. The residual area between the central spike and the bars can be interpreted as an upper bound measure of inequality attributed to differential efforts. The white labels at the bottom of each bar indicate the share of IOp in IO, i.e. the relative measure of inequality of opportunity IOR. The following circumstance sets are introduced sequentially: First (Sex), Second (Race), Third (Rural/Urban), Fourth (Average Family Income), Fifth (Educational Achievement Mother).

The results from Figure 1 show that heterogeneity across types is not negligible. The central spike indicates total outcome inequality as measured by the mean log deviation (MLD). As the sample is balanced on the last circumstance set, it is constant across the different scenarios. For the moment let's focus on the bars to the left of each spike, which show the MLD in Φ, i.e. the inequality in mean outcomes across types, for the linear (Equation (3)) and the non-linear (Equation (5)) case. The percentage figures at the bottom of each bar indicate the respective relative measures of IOp, IOR. Gradually introducing the circumstance sets, the difference between the linear and the non-linear case increases to more than 11 percentage points in the full-blown model. While the most extensive circumstance set for the linear case yields an IOR of 14.9%, the non-linear case yields a lower bound IOp measure of 26%. It is not surprising that the divergence between the two approaches follows a convex path as the introduction of each circumstance C^j adds $(x_j - 1) * \prod_i^{j-1} x_i$ new regressors to the estimation. Thus, the partition grows exponentially.[2] The exponential growth in parameters to be estimated serves as the main

[2] It is illustrative to compare the adjusted and the standard R^2-measures for the linear and the non-linear case, respectively (see Figure 2 in the Online Appendix). The penalization of the exponential growth in coefficients under the adjusted measure inflates the difference between the two statistics as we sequentially increase the number of circumstances under consideration. Yet it is apparent that the consideration of heterogeneity provides a strong upwards correction of the explained variance in the outcome variable even when considering the adjusted R^2-measure.

justification for relying on the linear parametric approach. As can be inferred from Table 5 in the Online Appendix, 16 coefficients are omitted from the full-blown model due to multi-collinearity. The median of observations per type T^k is 50, while 25% of all $\mu^k(p)$ are estimated using group sizes of <19. Thus, many coefficients may be measured very imprecisely. Furthermore, it is noteworthy that 14 interaction terms are significant at the 10%-level individually, however, a joint F-test on all interaction terms fails to reject the H_0 at the 10%-level. This contradictory finding may be attributed to the large number of highly collinear interaction terms that are introduced without yielding a corresponding decrease in the residual sum of squares. To be sure that the increase in the point estimate of IOR does not come at the cost of decreased precision we calculate confidence intervals for IOR using a bootstrap procedure with 1000 repetitions. In the largest circumstance set the 95% confidence band for the linear case is $[12.0\%, 17.7\%]$, the corresponding interval for the non-linear case is $[22.9\%, 29.0\%]$. We thus can conclude that the precision of the IOR estimate is not deteriorated by the introduction of the full battery of interaction terms.

We now turn to heterogeneity in effort levels. The two bars to the right of the central spike incorporate standardized effort distributions. The center-right bar presumes a homogeneous influence of circumstances, while the rightmost graph relaxes this assumption by incorporating both sources of heterogeneity (Equation (10)).

Table 2: Differentials across sources of heterogeneity

Circ. Set		Heterog. Circ.	Heterog. Effort	Both
First	Linear	0.0 (-)	1.2 (96.9)	1.2 (96.9)
First	Heterog. Circ.	- (-)	1.2 (96.9)	1.2 (96.9)
First	Heterog. Effort	- (-)	- (-)	0.0 (-)
Second	Linear	0.3 (38.1)	2.1 (81.1)	2.5 (92.5)
Second	Heterog. Circ.	- (-)	1.8 (62.9)	2.2 (83.0)
Second	Heterog. Effort	- (-)	- (-)	0.4 (40.3)
Third	Linear	0.8 (71.1)	2.3 (83.4)	3.1 (107.9)
Third	Heterog. Circ.	- (-)	1.4 (47.8)	2.3 (85.4)
Third	Heterog. Effort	- (-)	- (-)	0.9 (65.7)
Fourth	Linear	3.7 (145.8)	2.6 (77.2)	6.0 (164.5)
Fourth	Heterog. Circ.	- (-)	1.1 (27.0)	2.3 (81.4)
Fourth	Heterog. Effort	- (-)	- (-)	3.4 (130.1)
Fifth	Linear	11.2 (302.5)	3.3 (69.6)	13.0 (309.5)
Fifth	Heterog. Circ.	- (-)	7.9 (135.9)	1.9 (78.2)
Fifth	Heterog. Effort	- (-)	- (-)	9.8 (226.6)

Differences are given in percentage points. The associated t-values in parentheses are calculated using 1000 bootstrap repetitions. The following circumstance sets are introduced sequentially: First (Sex), Second (Race), Third (Rural/Urban), Fourth (Average Family Income), Fifth (Educational Achievement Mother).

We note that the standardization of type-specific effort distributions increases the point estimate of IOR by 3.2 percentage points in comparison with the standard estimation approach when maintaining the linearity assumption with respect to circumstances. The increase amounts to 1.9 percentage points when introducing circumstances non-linearly. Again the results are quite robust, with the 95-% confidence bands hovering between 7 and 9 percentage points. Table 2 summarizes the differences across the outlined approaches to heterogeneity for the different circumstance sets. It is noteworthy that for all differences, the H_0 of equality in IOR can be rejected at the 1%-level.

4 Conclusion

In this note, we have shown that neglecting type-specific heterogeneity in the influence of circumstances and efforts may have important implications for IOp estimates. It is well-known that the absence of data on all relevant circumstances renders estimates of IOp to be lower bounds (see Niehues and Peichl, 2014, for an upper-bound estimate). Unfortunately, applied researchers on IOp have little leverage to correct this shortcoming. However, even in the presence of data limitations IOp estimates could be considerably improved by taking type-specific heterogeneity seriously. Using the same limited circumstance set, we have increased the share of IOp in the observed outcome distribution by almost 50%. Furthermore, our calculations support the finding of Björklund et al. (2012) that type-specific variance in effort levels is another important determinant of IOp. Lastly, we want to highlight the importance of establishing the provision of standard errors as a good practice in applied works on IOp in order to afford a good sense for the precision of the results to the interested research community.

References

Balcázar, C. F. (2015). Lower bounds on inequality of opportunity and measurement error. *Economics Letters*.

Björklund, A., Jäntti, M., and Roemer, J. E. (2012). Equality of opportunity and the distribution of long-run income in Sweden. *Social Choice and Welfare*, 39(2-3):675–696.

Bourguignon, F., Ferreira, F. H. G., and Menéndez, M. (2007). Inequality of Opportunity in Brazil. *Review of Income and Wealth*, 53(4):585–618.

Checchi, D. and Peragine, V. (2010). Inequality of opportunity in Italy. *The Journal of Economic Inequality*, 8(4):429–450.

Ferreira, F. H. G. and Gignoux, J. (2011). The Measurement of Inequality of Opportunity: Theory and an Application to Latin America. *Review of Income and Wealth*, 57(4):622–657.

Ferreira, F. H. G. and Peragine, V. (2015). Equality of Opportunity: Theory and Evidence. *IZA Discussion Paper*, 8994.

Niehues, J. and Peichl, A. (2014). Upper bounds of inequality of opportunity: theory and evidence for Germany and the US. *Social Choice and Welfare*, 43(1):73–99.

Ramos, X. and Van de gaer, D. (2016). Empirical Approaches to Inequality of Opportunity: Principles, Measures, and Evidence. *Journal of Economic Surveys*, forthcoming.

Roemer, J. E. (1993). A Pragmatic Theory of Responsibility for the Egalitarian Planner. *Philosophy & Public Affairs*, 22(2):146–66.

Roemer, J. E. (1998). *Equality of Opportunity.* Harvard University Press, Cambridge.

Roemer, J. E. and Trannoy, A. (2015). Equality of Opportunity. *Journal of Economic Literature*, forthcoming.

5 Online Appendix

Table 3: Average Gross Income (25-30)

	(1)	(2)	(3)	(4)	(5)
Male	0.291*** (0.038)	0.281*** (0.038)	0.282*** (0.038)	0.287*** (0.037)	0.288*** (0.037)
Majority		0.356*** (0.038)	0.344*** (0.040)	0.146*** (0.042)	0.138** (0.042)
SMSA, Not center (16)			0.126** (0.048)	-0.020 (0.049)	-0.018 (0.048)
SMSA, Center (16)			0.026 (0.056)	-0.029 (0.055)	-0.029 (0.055)
Avrg. Net Fam. Inc. <P25				-0.708*** (0.064)	-0.576*** (0.069)
Avrg. Net Fam. Inc. <P50				-0.347*** (0.060)	-0.289*** (0.062)
Avrg. Net Fam. Inc. <P75				-0.129* (0.059)	-0.108 (0.060)
Secondary (16)					0.307*** (0.053)
Intermediate (16)					0.372*** (0.077)
College (16)					0.371*** (0.078)
Constant	9.574*** (0.027)	9.425*** (0.031)	9.359*** (0.049)	9.862*** (0.074)	9.540*** (0.090)
N	3149	3149	3149	3149	3149
F-Stat. |p-Value|					

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Average Gross Income (25-30)

	(1)	(2)	(3)	(4)	(5)
Male	0.291*** (0.038)	0.232*** (0.050)	0.432*** (0.120)	0.738 (0.847)	-1.966 (2.812)
Majority		0.300*** (0.054)	0.436*** (0.114)	0.548 (0.634)	-3.471 (2.861)
Male × Majority		0.113 (0.077)	-0.039 (0.161)	-0.766 (0.901)	3.923 (3.095)
SMSA, Not center (16)			0.250* (0.098)	-0.009 (0.612)	-5.879 (3.194)
SMSA, Center (16)			0.247* (0.101)	0.384 (0.653)	-1.720 (2.584)
Male × SMSA, Not center (16)			-0.146 (0.142)	-0.036 (0.868)	2.461 (2.616)
Male × SMSA, Center (16)			-0.348* (0.145)	-1.590 (0.931)	-0.238 (2.491)
Majority × SMSA, Not center (16)			-0.130 (0.136)	-0.057 (0.653)	3.193 (2.718)
Majority × SMSA, Center (16)			-0.267 (0.174)	-0.389 (0.732)	3.158 (2.613)
Male × Majority × SMSA, Not center (16)			0.075 (0.193)	0.368 (0.929)	-3.106 (2.826)
Male × Majority × SMSA, Center (16)			0.304 (0.249)	1.629 (1.049)	-0.642 (2.680)
Avrg. Net Fam. Inc. <P25				-0.563 (0.608)	-6.225 (3.280)
Avrg. Net Fam. Inc. <P50				-0.079 (0.621)	-6.404 (3.303)
Avrg. Net Fam. Inc. <P75				0.210 (0.642)	-4.652 (3.110)
Male × Avrg. Net Fam. Inc. <P25				-0.280 (0.860)	2.497 (2.822)
Male × Avrg. Net Fam. Inc. <P50				-0.416 (0.880)	3.011 (2.874)
Male × Avrg. Net Fam. Inc. <P75				-0.233 (0.912)	1.689 (2.405)
Majority × Avrg. Net Fam. Inc. <P25				-0.283 (0.665)	3.710 (2.879)
Majority × Avrg. Net Fam. Inc. <P50				-0.589 (0.668)	4.668 (2.914)
Majority × Avrg. Net Fam. Inc. <P75				-0.424 (0.687)	2.395 (2.619)
Male × Majority × Avrg. Net Fam. Inc. <P25				0.539 (0.949)	-4.638 (3.133)
Male × Majority × Avrg. Net Fam. Inc. <P50				1.092 (0.949)	-4.960 (3.183)
Male × Majority × Avrg. Net Fam. Inc. <P75				0.579 (0.977)	-3.139 (2.476)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25				0.241 (0.629)	6.039 (3.202)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50				0.030 (0.640)	6.595* (3.228)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75				-0.115 (0.663)	4.928 (3.006)

SMSA, Center (16) × Avrg. Net Fam. Inc. <P25	-0.205 (0.666)	1.709 (2.592)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P50	-0.255 (0.680)	3.071 (2.630)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P75	-0.280 (0.706)	1.282 (2.313)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25	-0.247 (0.892)	-2.714 (2.636)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50	-0.140 (0.909)	-3.338 (2.698)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75	-0.112 (0.942)	-1.517 (2.113)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25	1.239 (0.951)	-0.141 (2.509)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50	1.522 (0.971)	-1.174 (2.586)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75	0.987 (1.009)	0.652 (1.862)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25	-0.301 (0.717)	-3.458 (2.759)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50	0.105 (0.700)	-4.740 (2.794)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75	0.129 (0.717)	-1.925 (2.386)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25	0.373 (0.923)	-3.171 (2.734)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50	0.013 (0.801)	-5.779* (2.874)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75	0.104 (0.807)	-2.596 (2.169)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25	0.355 (1.020)	3.842 (2.916)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50	-0.457 (0.997)	5.143 (2.962)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75	-0.304 (1.021)	2.136 (1.984)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25	-1.496 (1.262)	1.609 (2.872)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50	-1.468 (1.149)	1.145 (3.074)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75	-0.939 (1.159)	-0.448 (1.434)
Secondary (16)		-5.493 (3.194)
Intermediate (16)		-5.441 (3.434)
College (16)		-4.038 (2.215)
Male × Secondary (16)		2.734 (2.616)
Male × Intermediate (16)		-2.174 (3.771)
Male × College (16)		1.243 (2.189)
Majority × Secondary (16)		4.152 (2.753)
Majority × Intermediate (16)		4.264 (3.210)

Majority × College (16)	2.390 (1.846)
Male × Majority × Secondary (16)	-5.047 (2.892)
Male × Majority × Intermediate (16)	0.593 (3.405)
Male × Majority × College (16)	-3.253 (2.728)
SMSA, Not center (16) × Secondary (16)	5.649 (3.115)
SMSA, Not center (16) × Intermediate (16)	6.002 (3.378)
SMSA, Not center (16) × College (16)	4.821* (2.080)
SMSA, Center (16) × Secondary (16)	2.037 (2.458)
SMSA, Center (16) × Intermediate (16)	2.715 (2.875)
SMSA, Center (16) × College (16)	0.558 (1.417)
Male × SMSA, Not center (16) × Secondary (16)	-2.359 (2.418)
Male × SMSA, Not center (16) × Intermediate (16)	1.785 (3.578)
Male × SMSA, Not center (16) × College (16)	-1.076 (1.901)
Male × SMSA, Center (16) × Secondary (16)	-0.822 (2.211)
Male × SMSA, Center (16) × Intermediate (16)	-1.499 (3.502)
Male × SMSA, Center (16) × College (16)	-0.849 (2.009)
Majority × SMSA, Not center (16) × Secondary (16)	-3.173 (2.613)
Majority × SMSA, Not center (16) × Intermediate (16)	-3.891 (3.120)
Majority × SMSA, Not center (16) × College (16)	-1.785 (1.593)
Majority × SMSA, Center (16) × Secondary (16)	-3.754 (2.451)
Majority × SMSA, Center (16) × Intermediate (16)	-4.188 (3.246)
Majority × SMSA, Center (16) × College (16)	-1.543 (1.770)
Male × Majority × SMSA, Not center (16) × Secondary (16)	3.635 (2.620)
Male × Majority × SMSA, Not center (16) × Intermediate (16)	-0.688 (3.084)
Male × Majority × SMSA, Not center (16) × College (16)	1.833 (2.383)
Male × Majority × SMSA, Center (16) × Secondary (16)	2.238 (2.339)
Male × Majority × SMSA, Center (16) × Intermediate (16)	-0.546 (2.252)
Male × Majority × SMSA, Center (16) × College (16)	1.899 (2.642)
Avrg. Net Fam. Inc. <P25 × Secondary (16)	5.833 (3.202)

Avrg. Net Fam. Inc. <P25 × Intermediate (16)	5.729 (3.464)
Avrg. Net Fam. Inc. <P25 × College (16)	3.843 (2.280)
Avrg. Net Fam. Inc. <P50 × Secondary (16)	6.497* (3.227)
Avrg. Net Fam. Inc. <P50 × Intermediate (16)	7.398* (3.498)
Avrg. Net Fam. Inc. <P50 × College (16)	3.662 (2.332)
Avrg. Net Fam. Inc. <P75 × Secondary (16)	4.652 (3.010)
Avrg. Net Fam. Inc. <P75 × Intermediate (16)	5.287 (3.316)
Avrg. Net Fam. Inc. <P75 × College (16)	3.676* (1.821)
Male × Avrg. Net Fam. Inc. <P25 × Secondary (16)	-2.898 (2.635)
Male × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	2.352 (3.705)
Male × Avrg. Net Fam. Inc. <P25 × College (16)	-0.923 (2.484)
Male × Avrg. Net Fam. Inc. <P50 × Secondary (16)	-3.592 (2.699)
Male × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	0.452 (3.880)
Male × Avrg. Net Fam. Inc. <P50 × College (16)	-0.792 (2.400)
Male × Avrg. Net Fam. Inc. <P75 × Secondary (16)	-1.705 (2.138)
Male × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	1.314 (2.138)
Male × Avrg. Net Fam. Inc. <P75 × College (16)	-0.774 (1.268)
Majority × Avrg. Net Fam. Inc. <P25 × Secondary (16)	-4.174 (2.785)
Majority × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	-3.925 (3.334)
Majority × Avrg. Net Fam. Inc. <P25 × College (16)	0.909 (1.569)
Majority × Avrg. Net Fam. Inc. <P50 × Secondary (16)	-5.712* (2.820)
Majority × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	-5.568 (3.331)
Majority × Avrg. Net Fam. Inc. <P50 × College (16)	-2.774 (2.082)
Majority × Avrg. Net Fam. Inc. <P75 × Secondary (16)	-2.803 (2.479)
Majority × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	-3.141 (3.056)
Majority × Avrg. Net Fam. Inc. <P75 × College (16)	-1.160 (1.172)
Male × Majority × Avrg. Net Fam. Inc. <P25 × Secondary (16)	5.837* (2.958)
Male × Majority × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	-0.489 (3.139)
Male × Majority × Avrg. Net Fam. Inc. <P25 × College (16)	0.000 (.)

Male × Majority × Avrg. Net Fam. Inc. <P50 × Secondary (16)	6.808* (3.009)
Male × Majority × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	0.534 (3.620)
Male × Majority × Avrg. Net Fam. Inc. <P50 × College (16)	3.866 (2.996)
Male × Majority × Avrg. Net Fam. Inc. <P75 × Secondary (16)	3.881 (2.173)
Male × Majority × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	-0.179 (1.738)
Male × Majority × Avrg. Net Fam. Inc. <P75 × College (16)	2.399 (1.635)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)	-5.559 (3.129)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	-6.042 (3.428)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × College (16)	-3.269 (2.389)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)	-6.521* (3.157)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	-7.477* (3.461)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × College (16)	-3.977 (2.252)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)	-4.644 (2.910)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	-5.474 (3.260)
SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × College (16)	-3.960* (1.598)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)	-1.720 (2.473)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	-2.122 (2.959)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × College (16)	0.810 (2.047)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)	-3.375 (2.516)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	-5.438 (2.975)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × College (16)	-1.081 (1.659)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)	-1.219 (2.146)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	-2.930 (2.700)
SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × College (16)	0.000 (.)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)	2.286 (2.456)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	-1.704 (3.546)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × College (16)	0.593 (2.677)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)	3.225 (2.530)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	-0.186 (3.723)

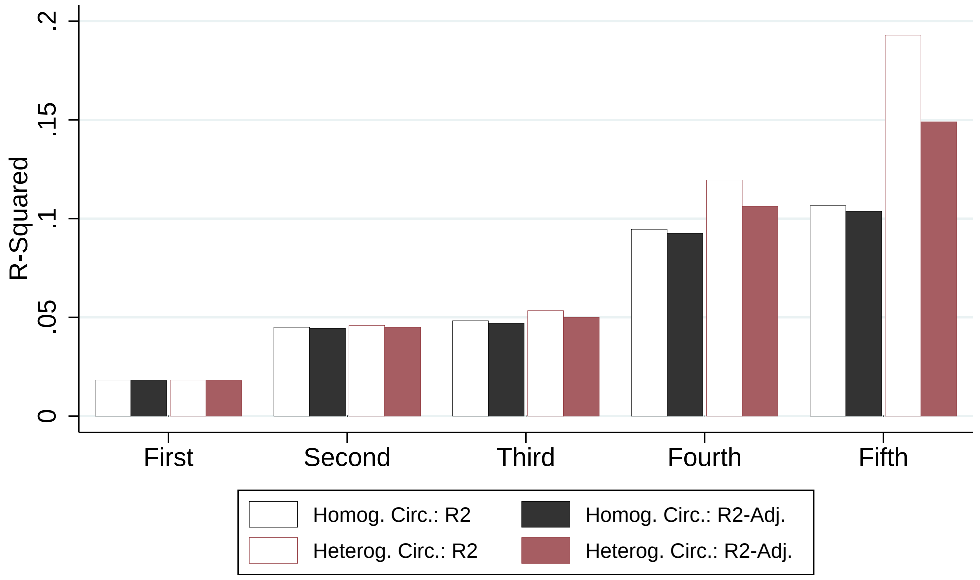
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × College (16)	0.176 (2.235)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)	1.041 (1.829)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	-1.291 (1.637)
Male × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × College (16)	0.000 (.)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)	0.920 (2.246)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	0.780 (3.369)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × College (16)	0.000 (.)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)	2.270 (2.342)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	4.427 (3.659)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × College (16)	1.704 (2.324)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)	-0.596 (1.391)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	2.578 (2.578)
Male × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × College (16)	0.000 (.)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)	3.068 (2.684)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	3.853 (4.287)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × College (16)	0.000 (.)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)	5.204 (2.709)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	5.230 (3.285)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × College (16)	2.584 (1.991)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)	1.913 (2.247)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	2.744 (2.930)
Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × College (16)	0.000 (.)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)	3.387 (2.793)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)	4.295 (3.636)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)	6.298* (2.755)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × Intermediate (16)	6.993 (3.576)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × College (16)	3.457 (2.529)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)	2.763 (1.927)
Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)	4.630 (3.064)

Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × College (16)					0.000 (.)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)					-3.798 (2.771)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)					0.000 (.)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P25 × College (16)					0.000 (.)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)					-6.268* (2.802)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × Intermediate (16)					-1.808 (3.411)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P50 × College (16)					-3.069 (2.873)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)					-2.464 (1.634)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)					0.000 (.)
Male × Majority × SMSA, Not center (16) × Avrg. Net Fam. Inc. <P75 × College (16)					0.000 (.)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Secondary (16)					-3.297 (2.835)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P25 × Intermediate (16)					0.000 (.)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × Secondary (16)					-2.629 (2.840)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P50 × College (16)					-2.033 (3.439)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Secondary (16)					0.000 (.)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × Intermediate (16)					0.000 (.)
Male × Majority × SMSA, Center (16) × Avrg. Net Fam. Inc. <P75 × College (16)					0.000 (.)
Constant	9.574*** (0.027)	9.449*** (0.035)	9.244*** (0.083)	9.581*** (0.599)	15.057*** (3.276)
N	3149	3149	3149	3149	3149
F-Stat. $\lvert$p-Value$\rvert$		2.159 $\lvert$0.142$\rvert$	1.478 $\lvert$0.170$\rvert$	1.265 $\lvert$0.124$\rvert$	1.125 $\lvert$0.139$\rvert$

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2: (Adj.) R-Squared: Linear vs. non-linear case



Note: The hollow bars yield the standard R-squared, while the solid bars indicate the value of the adjusted R-squared measure. The black bars indicate yield the two measures for the linear model respectively, whereas the bars in maroon are indicative for the fully interacted model. The following circumstance sets are introduced sequentially: First (Sex), Second (Race), Third (Rural/Urban), Fourth (Average Family Income), Fifth (Educational Achievement Mother).