The Trade-off between Child Quantity and Quality

Eric A. Hanushek

University of Rochester

An empirical investigation of trade-offs between number of children and their scholastic performance confirms that family size directly affects children's achievement. Though parents show no favoritism to first-born children, being early in the birth order implies a distinct advantage, entirely because of the higher probability of being in a small family. Recent large changes in family size explain a portion of aggregate test score declines, but increased divorce rates and market work by mothers have no apparent impact. Finally, teachers are shown to differ enormously, even though performance differences are poorly captured by commonly measured teacher characteristics. The evidence supports a teacher skill interpretation of differences in classroom achievement.

Economists have increasingly turned their attention to behavior within families because of its direct implications for such diverse issues as population growth, intergenerational transfers of wealth, human capital accumulation, and macroeconomic policy. While this analysis can be related to each of these issues, it is easiest to motivate it in terms of modern economic demography, which concentrates on parental trade-offs between the number and quality of children. This theory, advanced by Becker (1960), Becker and Lewis (1973), and Willis (1973), explains among other things how birth rates could fall with increasing income even though children are not inferior goods.

Helpful comments were provided by Stanley Engerman and seminar participants at Princeton, Williams, and Wisconsin. This work was partially supported by the National Institute of Education. The data were provided by Mathematica Policy Research, and this research could not have been conducted without its help.

[*Journal of Political Economy*, 1992, vol. 100, no. 1]

However, direct evidence on this hypothesized trade-off is scant, owing to a lack of suitable data on child quality, to considerable ignorance about the "production function" for child quality, and to an inability to separate confounding factors including other family attributes and exogenous school inputs. This paper exploits an exceptionally rich body of data to develop empirical evidence on the magnitude of quantity-quality trade-offs and, as a by-product, gives new insights into the interaction between families and schools and into variations in school and teacher quality.

Most previous studies of the effects of family circumstances on children's performance have been very narrowly focused. While both demographers and psychologists have long been interested in the effects of family size, birth order, and child spacing on scholastic performance, only passing attention has typically been given to other factors that might affect scholastic achievement. For example, as reviewed in Blake (1989), income measures are frequently omitted from birth order and family composition studies, even though most theoretical and empirical analyses of fertility decisions suggest that family size is related to income. A related policy perspective, sparked by the rise in female labor force participation and in the incidence of one-parent families, has focused on the potential long-run impacts of family circumstances on the transmission of human capital (see, e.g., Congressional Budget Office 1987). The supporting empirical research has, however, also been clouded by many specification issues, including the near total neglect of how schools affect children's achievement.

This study follows a different tradition—that of educational production studies. The investigation of family composition is embedded in a more general model of students' scholastic achievement, the measure of child quality employed here. It is important to note that because school choices are related to family circumstances, ignoring school inputs could lead to systematically overstating the importance of family factors for children's achievement. Data from the Gary Income Maintenance Experiment support a unique investigation of how specific family and school factors combine over time to determine a student's performance.

I. Conceptual Aspects of Family Effects

An important theoretical innovation in considering family behavior is the introduction of a production function for home activities. In the context of fertility decisions, families are seen as maximizing utility (which has arguments of the quantity and quality of children along with other goods) subject to the production function for child quality, a budget constraint, and a time constraint. The trade-off between child quantity and quality enters essentially because parents' time and resources must be spread thinner with more children. Such initial investigations of family behavior have been pushed considerably farther by work on the intergenerational transmission of wealth (see Becker and Tomes 1976, 1979; Behrman, Pollak, and Taubman 1982, 1986, 1989; Behrman and Taubman 1986). While differing in detail, these investigations help clarify the role of preferences and opportunities in determining income distribution and family well-being. These analyses of family behavior, motivated largely by issues of financial transactions within households, depart from the original research, which emphasized the role of time budgets. The analysis here, however, returns to the original focus on family time allocation.

The underlying conceptual model is a straightforward family maximization model in which parents choose time allocations to maximize an objective function defined in terms of total (or average) achievement by their children. This rather standard maximization problem, in which choices are constrained by the production function for student achievement and by available time, is extended in three ways: (1) by considering the dynamics of time allocation, (2) by allowing for the heterogeneous nature of time, and (3) by incorporating alternative perspectives on intrafamily equity.

Explicit consideration of how parental decisions evolve has important implications for the specification and interpretation of family behavioral models. The starting point is a particularly simple formulation of the dynamics of time allocation: At each instant, parents are assumed to make time allocations based on the then-existent number of children. Even this simple formulation of time dynamics, however, shows the importance of explicit consideration of family circumstances. For example, because each sibling faces a different effective path of family sizes, children in the same family do not end up at the same point, other things equal. Further, depending on child spacing, the same completed family size can correspond to widely varying parental time inputs to children during preschool and schooling periods.

Time allocation decisions must also consider that time is not homogeneous. The conceptual work here considers parental optimizing decisions about allocations of two types of educational inputs: "public" time and "private" time. The concept of public time, alluded to earlier by Hill and Stafford (1974), is directly analogous to a pure public good: all children share in this, and "consumption" of public time by one child does not lower the amount available to other children. Public time corresponds to group activities within the family: setting the general educational tenor of the family, motivating children toward schooling and careers, transmitting language, and so forth. Private time is child specific: activities with no spillovers to other children in the family such as helping with homework. It is natural to think that private time has greater educational value but is more expensive, because private time for any child subtracts from the total time available to other children. Such considerations highlight the substitution possibilities available to parents with varying numbers of children.

Finally, three easily derived alternatives illustrate the potential implications of parental views about achievement differences within the family.

1. If parents employ *nondiscriminatory time allocation* such that private time is equalized across children, the number of children acts simply as the "price" of private time. As the number of children increases, parents substitute public for private time, and in general the achievement of each child will fall with larger families but the decline will be less than linear.

2. If parents pursue *achievement maximization* and do not constrain themselves to equal time allocations, parents will tend to reinforce prior achievement differences in their children (as long as parental time is complementary with individual child ability).[1]

3. If parents pursue *compensatory time allocation* such that they are concerned with both the level and dispersion of family achievement, the distribution of private time across children is governed by two opposing forces. Maximizing average achievement dictates (with complementarity) spending more time with the more able (the "efficiency" effect), whereas minimizing the variance dictates taking time away from the most able and giving it to the least able (the "wealth" effect). While the precise details depend on the form of the household's objective function, private time will normally be devoted in greater amounts to the less able than would be the case with achievement maximization. All other things equal, however, the amount of compensatory adjustment will be less in larger families.

In all cases, achievement of each child will fall with the addition of more children. However, in the first case the achievement of any

[1] Becker and Tomes (1976) essentially conclude that this type of behavior is most likely. Note that families may still be "egalitarian" with respect to their children. Families can maximize aggregate welfare of children by giving time (education) to those who most benefit in terms of achievement and money (perhaps through bequests) to those who benefit less from time. Alternatively, parents may be able to "force" nonaltruistic children to act in an altruistic manner toward siblings (cf. Becker 1991). If parents perceive limits on their ability to transfer income (which may be particularly relevant in lower-income families), compensatory allocations (below) can also be consistent with the "child-neutral" case of Becker and Tomes (1976). Achievement maximization is a special case of "reinforcing strategy" as described in Behrman et al. (1982).

particular child will not depend on the distribution of achievement of other children in the family; in the second case achievement will be higher with both higher own ability and lower relative ability of siblings; and in the last case an individual who is smarter than his or her siblings may receive less attention than in the second case, with the amount depending on the distaste for intrafamily differences and the magnitude of within-family differences.

In this structure, attention focuses on parental time allocations in any period. Parents are assumed to treat all children "evenhandedly": no special attention is given to the first born or to the "baby" of the family. But, as mentioned, this does not imply achievement equality. Educational inputs accumulate so that, independent of ability differences or different profiles of school inputs, children in different birth order positions will have different amounts of family inputs simply because of the sequential nature of births. Similarly, while the age distribution of children does not enter into the optimization at any point in time, different child spacings will imply different accumulated family inputs. It is important to note that the structure of the model with respect to each of these aspects is testable. The implications of each are discussed in Section IV.

II. Empirical Implementation

While the actual time allocations, particularly when private time and public time are distinguished, are difficult to observe directly, their underlying determinants can be observed and the reduced-form relationships can be estimated. The empirical analysis utilizes data about family characteristics, parental work behavior, incomes, and so forth that were generated over a 4-year period by the Gary Income Maintenance Experiment. These data were merged with information about achievement and school experiences of children from the experimental families between 1971 and 1975. The student schooling information included data on specific teachers along with test scores from the Iowa Reading Comprehension and Vocabulary tests. Some families received payments under alternative negative income tax schemes, and others were part of the experimental control group. All families had relatively low incomes: the average income of sampled families (in 1974 dollars) was approximately $6,500, somewhat above the 1974 poverty line for an urban family of four, which stood at $5,038. All families were black. (Descriptive statistics are found in App. table A1.)

The theoretical model, which focused on the "instantaneous" flow of achievement, has direct implications for the specification of the statistical models. The optimal allocation of time will change with variations in exogenous factors, with variations in family size and structure, and so forth. Thus in terms of the level of achievement, the entire past history, including the time path of inputs and family decisions, will be relevant. Furthermore, because the short-run and long-run impacts of both family and nonfamily factors may differ significantly, accurate identification of the time profile specific to each individual is essential. A corollary of this is that aggregation of inputs across time, which is frequently done by relying exclusively on contemporaneous values of inputs in a cross-sectional analysis of achievement, creates a classic errors-in-variables problem (Hanushek 1979).

Past consideration of measurement and estimation problems, while focusing on the contribution of school inputs, has demonstrated that analysis of "value added," or achievement growth, over a restricted period of time can be used to circumvent the most serious difficulties. In particular, if one ignores for the moment any randomness, the achievement of the ith student at time t (A_{it}) can be thought of as a function of the cumulative inputs of families ($F_i^{(t)}$) and schools ($S_i^{(t)}$) along with the cumulative inputs of other exogenous factors ($X_i^{(t)}$). The level of achievement at any point can be written simply as

$$A_{it} = \Phi(F_i^{(t)}, S_i^{(t)}, X_i^{(t)}). \tag{1}$$

If achievement at a previous time (t^*) is also observed, it is possible to concentrate on value added over the intervening period, such as

$$A_{it} = \phi(F_i^{(t-t^*)}, S_i^{(t-t^*)}, X_i^{(t-t^*)}, A_{it^*}). \tag{2}$$

In this formulation, the entire past history of inputs is not needed; instead, just data on inputs over the limited interval ($t - t^*$) are needed. This also matches the previously discussed family time allocation problem; there, it was assumed that family size and structure were constant for the decision period.[2] This general formulation of the "value-added" specification lessens the data requirements, but it does so at the expense of some additional assumptions about the relationships (see, e.g., Hanushek 1979, 1986; Ragosa and Willett 1985). The value-added formulation also accounts for any fixed, but unmeasured, effects such as differences in innate ability, motivation, and so forth as long as they have a proportional effect on achievement (in the logarithmic models estimated below).

Empirical implementation, which utilizes data on inter- as well as intrafamily variations in achievement, must consider three possible modifications to the basic theoretical structure: (1) the amount of

[2] Note that direct measures of the family's public and private time inputs are not included. These are "solved out," and the relationships are estimated in terms of the underlying determinants of time allocations. The alternative of including direct time estimates has been pursued by others (e.g., Stafford 1987), but this is subject to a variety of difficult measurement concerns.

total time to be allocated may differ across families, (2) the quality of the time allocated to children may also differ across families, and (3) other exogenous factors almost certainly influence children's achievement.

Two major systematic differences across families in total time available are considered: the presence or absence of a father and the work behavior of the parents. Analyses of these differences also provide information about the potential long-run effects of increased divorce rates and the dramatic changes in female labor force participation rates of the past two decades.

While the quality of parental time is difficult to observe directly, several measures of family socioeconomic status provide proxies. Since the quality of time as related to child-rearing practices would be expected to evolve slowly over time, long-run measures of the socioeconomic status of the family seem most appropriate, and two alternative measures are used: "permanent" income and parental education levels. Distinctions between permanent and current family circumstances have not consistently been made in past work, but this distinction is important from both an analytical and a policy view. Contemporaneous measures, such as current income, not only will be noisy measures of fundamental quality differences but also will confuse current conditions—such as changes in labor force behavior or changes in the number of parents—with longer-run quality considerations.

The difference between quantity and quality of parental time is somewhat blurred. Direct analyses of reported time allocations, in particular the work of Hill and Stafford (1974, 1980), suggest that quantity of time may vary directly with socioeconomic status (see also Leibowitz 1974; Murnane, Maynard, and Ohls 1981). Further, income can clearly enter directly into the achievement relationships through material advantages or purchased inputs. For this analysis, however, in which quality of time is assumed to augment the amount of available time, precise differentiations are not important.

The most important exogenous factors to be considered relate to teachers and school inputs. While a variety of approaches to the measurement of school resources have been taken in past work, a particularly simple and general formulation of school effects is employed in the main analysis. Teacher quality, or "skill," is viewed as being idiosyncratic, differing across teachers but not in accord with simply measurable attributes. Skill differences are estimated by mean differences in student achievement growth across teachers, conditional on other inputs.[3] When this general covariance structure, which is equiv-

[3] Separate teacher, or classroom, effects are estimated only if the teacher is observed with three or more sampled students.

alent to including separate dummy variables for each teacher in the sample, is used, there is no requirement to specify or measure the precise characteristics of teachers and schools that are important—a task that has proved extremely difficult (see Hanushek [1986] for an extended discussion). A subsequent section investigates the possibility that differences in classroom performance are not entirely attributable to differences in teacher skill but instead involve more complicated interactions of teachers and specific classes of students. This section also considers directly the relationship between achievement and a set of commonly measured characteristics of teachers.

III. Empirical Results

The estimation is divided into two parts: preschool and school. For the preschool period, for which no measures of prior achievement (A_{it*}) are available, value-added models cannot be estimated, and specifications like equation (1) are employed. For the school period, value-added models are consistently estimated.

A. *Achievement Growth during Schooling*

Achievement growth for grades 2–6 are specified in value-added form like equation (2). The individual grade-level samples and different school years are pooled, although separate intercepts and coefficients on beginning achievement are estimated for each grade. All the estimates still refer to change in achievement across a single grade level (e.g., achievement growth from the end of the fourth through the end of the fifth grades). The estimation is done in log-log form so that the parameters have the usual elasticity interpretation. The distribution across grades and calendar years of the 1,920 students for whom complete data are available is displayed in Appendix table A1.

Because of the possibility of measurement errors on the tests, the models are estimated using maximum likelihood methods that allow for the error variance in the prior achievement measures. In this, the prior achievement level is viewed as an error-prone measure of "true" ability such as

$$M_{it*} = A_{it*} + \epsilon_{it*}, \tag{3}$$

where M_{it*} is measured prior ability, which includes error, ϵ_{it*}. For the estimation, ϵ_{it*} is assumed to have mean zero with a constant variance (σ^2) and to be independent of the true ability (A_{it*}). The published reliability coefficients for the specific test and grade level (which range from .86 to .94) are used to estimate the error variances (σ^2).[4]

Table 1 provides variable definitions, and table 2 displays the estimation results. The first two models for each test (cols. 1 and 2 for vocabulary achievement and cols. 4 and 5 for reading achievement) include dummy variables for individual teachers, and the last models (cols. 3 and 6) ignore differences in school inputs.

Teacher differences have dramatic effects on student performance within this sample (the nature of these differences is discussed in Sec. V). It is important to note here, however, that inclusion of teacher effects has little systematic effect on the remaining coefficients. This finding of little bias from omitting school inputs, however, could well be an artifact of the specific sample. All children are drawn from the same school system, thus limiting the variation in school inputs and the range of correlation between family and school characteristics.

The systematic effects of families are portrayed in the top portion of the table. In the log-log formulations that include entering levels of achievement (PRE variables for each grade), the coefficients indicate how annual achievement growth is affected by family inputs.

The results confirm a fundamental trade-off between quantity and quality of children. The elasticity with respect to number of children (KIDS) is approximately −.03 for both achievement measures. These estimates imply that annual achievement growth of each child in a family will fall by about 2 percent when a second child is added and about 0.5 percent when a sixth child is added.[5] The achievement of each child in the family declines as family size increases, but the rate of decline lessens, presumably reflecting substitution of parental education from individual-specific, or private, time to more public time.

Further confirmation of the basic model emphasizing aggregate competition for parental time comes from additional tests. Specific position in the birth order (i.e., being first born or last born) provided no additional information in explaining achievement growth. Neither did the age distribution of children as measured by the presence and

[4] The reliability coefficients should be interpreted as providing a lower bound on the error variance since they assume no systematic individual component to the errors. Further, there is no information on the "external validity" of the tests. Reliability coefficients for the different grade levels are given in App. table A1. Other investigations of the importance of measurement error have yielded mixed results. Harbison and Hanushek (1992) find that alternative corrections employing instrumental variables have little impact on estimated models of student achievement. That analysis demonstrates that there are no statistical difficulties with using prior achievement on the right-hand side. A general discussion of change in measured test scores is found in Cronbach and Furby (1970).

[5] Separate estimates were obtained by a nonparametric approach, and the estimated achievement drop-off followed the same pattern as the parametric estimates. The size-specific estimates were, however, more imprecise than those presented.

TABLE 1

VARIABLE DEFINITIONS

	Dependent Variables
READING*	Log of Iowa reading comprehension test (raw scores)
VOCABULARY*	Log of Iowa vocabulary test (raw scores)
	Independent Variables
PRE(*X*)*	Log of Iowa reading comprehension or vocabulary test (raw scores), grade *X*
GRADE *X*	Equals one for grade *X* and zero for all other grades
PERM*	"Permanent" income (average income, 1970–75)
FALL	Equals one if pretest taken in fall, zero if taken in spring
FEMALE	Equals one if girl, zero if boy
KIDS*	Number of children
FAMTEST	Equals one if test scores available for other members of family, zero if not available
RELFAM*	Performance of siblings relative to specific child (equals zero if FAMTEST equals zero)
MALE PRESENT	Equals one if adult male present, zero otherwise
MALE CHANGE	Equals one if adult male moved in during year; minus one if adult male moved out during year; zero if no change in adult male in household
MOTHER WORKS	Equals one if mother worked during year, zero otherwise
MOTHER'S HOURS	Average weekly work hours: TOTAL, ACADEMIC year, and SUMMER
WORKS LOT	Equals one if mother's work hours are greater than 30 hours per week, zero otherwise
AVER KIDS*	Average number of children in family, birth to grade 1
YOUNGER	Equals one if younger siblings present, zero otherwise
NUMYNG*	Number of younger siblings present

* Denotes logarithms of variables.

number of preschool children nor the average spacing of children within each family. In sum, the number of children completely characterizes the effects of family structure on achievement growth over each interval.

Direct estimates of within-family performance differences, which are uniquely possible within this data set, indicate that parents are not behaving like achievement maximizers. The variable RELFAM measures variations in performance *within families*. By comparing the average performance of siblings relative to the initial performance of the specific child, one can separate differential treatment of siblings from the absolute level of achievement across students and families.[6]

[6] Performance of siblings across tests and grade levels is averaged in terms of standard deviations away from the sample mean. These are then converted to grade-level

TABLE 2

ACHIEVEMENT GROWTH MODELS (Grades 2–6)

Variable	Vocabulary* (1)	Vocabulary* (2)	Vocabulary* (3)	Reading* (4)	Reading* (5)	Reading* (6)
KIDS*	−.028	−.038	−.027	−.025	−.030	−.029
	(−2.3)	(−3.5)	(−2.1)	(−2.1)	(−2.9)	(−2.3)
PERM*	.035	.037	.029	.017	.028	.028
	(2.8)	(3.0)	(2.2)	(1.4)	(2.3)	(2.3)
FEMALE	.002	.000	.017	.028	.041	.041
	(.2)	(.1)	(1.4)	(2.6)	(3.7)	(3.7)
FAMTEST	−.017	...	−.001	−.011	...	−.006
	(−1.2)		(−.1)	(−.8)		(−.4)
RELFAM*	.075	...	.108	.013	...	.031
	(2.5)		(3.3)	(.4)		(1.0)
FALL	−.213	−.215	−.210	−.178	−.177	−.161
	(−6.6)	(−6.7)	(−7.6)	(−5.5)	(−3.5)	(−6.1)
			Grade Level and Entering Achievement			
GRADE 2	.296	.319	.228	.244	.248	.185
	(3.6)	(3.9)	(2.8)	(2.9)	(3.0)	(2.2)

GRADE 3	.377	.351	.365	.138	.245	−.128
	(4.1)	(4.3)	(4.5)	(1.7)	(3.0)	(1.6)
GRADE 4	−.157	−.165	−.042	−.058	−.142	.161
	(−2.0)	(−2.0)	(−.5)	(−.7)	(−1.7)	(2.0)
GRADE 5	−.049	−.056	−.162	.029	−.059	.065
	(−.6)	(−.7)	(−2.0)	(.3)	(−.7)	(−.8)
PRE2*	.675	.642	.636	.497	.491	.508
	(9.4)	(9.1)	(8.5)	(7.4)	(7.4)	(7.4)
PRE3*	.610	.582	.707	.517	.513	.596
	(12.1)	(11.8)	(13.6)	(10.5)	(10.7)	(12.3)
PRE4*	.536	.504	.575	.444	.438	.381
	(11.7)	(11.5)	(10.9)	(10.7)	(11.0)	(9.2)
PRE5*	.505	.473	.624	.441	.436	.490
	(11.3)	(11.0)	(14.2)	(10.0)	(10.3)	(11.7)
PRE6*	.548	.516	.591	.556	.550	.595
	(13.7)	(13.5)	(14.2)	(12.1)	(12.5)	(12.6)
Intercept	.574	.606	.601	.690	.693	.529
	(4.7)	(5.0)	(4.9)	(5.6)	(5.7)	(4.5)
TEACHERS	yes	yes	no	yes	yes	no
Degrees of freedom	1,760	1,762	1,902	1,760	1,762	1,902
R^2	.66	.66	.50	.65	.65	.55

* See table 1.

For the vocabulary model (col. 1), the positive and statistically significant relationship indicates that, everything else equal, an individual with brighter siblings has higher achievement. In terms of the alternative models of parental behavior, this suggests compensatory time allocations, a result that is consistent with the theoretical predictions of Becker and Tomes (1979).[7] With reading achievement, this intrafamily relationship is less apparent: while still positive, the magnitude of within-family effects is much smaller, and the coefficient is no longer statistically significant. A lack of relationship between intrafamily differences and an individual child's achievement is consistent with nondiscriminatory time allocations. Thus in neither case is there any support for achievement maximization behavior of the parents.

Permanent income (PERM), included to index the educational environment of the home and the quality of parental time, affects achievement systematically. Measures of parental education add no information after permanent income is considered. Further, current income (in the year studied) was insignificantly related to achievement growth, reinforcing the interpretation that this is an index of underlying educational quality of the home and not specific material advantages. Thus it appears that the naive policy option of simply adding money may have little effect unless there is a concomitant change in parental behavior and educational interactions.

Girls outperform boys in reading growth, and the differences in vocabulary performance are small and insignificant. The advantage in reading scores is 3–4 percent.

The data permit indirect estimates of achievement growth over the summer, and there is no indication of a summer falloff. Earlier researchers suggest that much of the difference in test performance between blacks and whites can be attributed to differential learning over the summer months, with blacks losing much of what they gained in the prior school year (Hayes and Grether 1969; Heyns 1978). With the exception of one school year, the sampled pretest

---

equivalents for the grade of the specific child under consideration, and the performance of siblings relative to each student's initial achievement is calculated. This variable is set equal to zero if no test information on siblings is available (because either the student is an only child or siblings were not tested). No adjustments are made here for test reliability, although tests for siblings are averaged and thus should have a smaller error variance. The dummy variable FAMTEST indicates whether or not relative test scores are available. There is no prior on the sign or the magnitude of this effect, and it is uniformly insignificantly different from zero. If RELFAM were available for all sampled students, it could be interpreted as simply a measure of overall family levels instead of the relative position, as indicated here.

[7] An alternative explanation frequently given is that siblings contribute to education (cf. Zajonc and Markus 1975; Zajonc 1976). In such a case, however, the absolute level of siblings' achievement, instead of the relative performance, would be the relevant factor.

TABLE 3

MARGINAL EFFECTS OF PARENTAL INPUT AND WORK BEHAVIOR

	VOCABULARY*		READING*	
VARIABLE	(1)	(2)	(1)	(2)
MALE PRESENT	−.0027 (−.2)	−.0017 (−.1)	.0017 (.1)	.0027 (.2)
MALE CHANGE		−.0051 (−.2)		.0130 (.6)
MOTHER WORKS	−.0208 (−.6)	−.0326 (−.7)	.0036 (.1)	−.0026 (−.1)
MOTHER'S TOTAL HOURS	.0005 (.5)		.0002 (.2)	
MOTHER'S ACADEMIC HOURS		.0013 (.4)		.0006 (.2)
MOTHER'S SUMMER HOURS		.0004 (.0)		.0003 (.4)
MOTHER'S HOURS > 30		−.0008 (−.3)		−.0005 (−.2)

NOTE.—The t-statistics for each coefficient are found in parentheses. Estimated models include all other variables shown in table 2, col. 1 or 4.
* See table 1.

scores represent achievement of the student at the end of the previous grade. In the first year of the data, however, the pretest of the sixth grade came in October of that year instead of May of the previous academic year. The dummy variable (FALL) identifying this situation indicates that achievement grows 16–21 percent over the months between tests.[8] Thus, in this sample of low-income blacks, no summer losses are found, and such explanations of racial differences in achievement are not supported.

Regardless of how it was specified, work behavior of mothers has no apparent effect on children's achievement. Table 3, which extracts just the relevant coefficients, indicates that neither the act of working nor differential amounts of time spent in the labor force systematically affect children's achievement. A variety of alternative specifications of mother's work, including investigating different effects of work during the academic year and summer and different nonlinear

[8] The exact interpretation of these gains depends on assumptions about the pace of learning during the school year, since the tests were given in May and in October. If one presumes that not much learning goes on during June and September, these estimates indicate noticeable learning over the summer months. If, however, growth through the year is roughly linear from September through June, one would expect June and September to contribute 15–20 percent to performance, implying no learning in the summer. There is in either case no evidence of the falloff in achievement suggested by some (see the review by Heyns [1978]). Specific year dummies for the remaining years were consistently insignificant, reinforcing the view that this is a time of testing effect and not a specific year effect.

effects, were investigated, but the estimated coefficients were never greater than their standard errors. *Negative* effects on children's learning do not appear to be a problem, a very important finding for assessing the potential for long-run effects of increased female labor force participation. This is consistent with the general findings in other studies, which find no systematic positive or negative effect of mother's work behavior on children's development (Kamerman and Hayes 1982; Hayes and Kamerman 1983). While data on the kind or quality of any child care services used are unavailable, the services chosen do not systematically deter children's achievement growth.

Somewhat more surprising is the finding, also displayed in table 3, that the presence or absence of an adult male in the family appears to have no effect, at least when income is held constant. This appears true in both the short and the long run. During the year of any family change (from a two-adult to a one-adult family or vice versa), there is no systematic impact on the scholastic performance of children. Nor is there a systematic difference in achievement growth between families with one and families with two parents throughout any given year. These findings, which differ from other studies on the achievement effects of single-parent families, may simply reflect the fact that qualitative aspects of any family split or creation dominate or may reflect inadequate control for differences in income and race in previous studies (see Congressional Budget Office 1987). Finally, changing schools per se does not systematically affect children. It matters considerably which teacher a student has, and therefore changes in schools could ultimately have a large effect. With the specific teachers and school resources at the classroom level held constant, however, the change of schools within the system has no independent effect on achievement.

Each of these findings tells a consistent story: Children are quite resilient to changes in their environment, and, presumably, a variety of things such as divorce or a change of schools can be compensated for so that there is no systematic effect.

B. Preschool Achievement

The preschool models can be viewed as models of "entering" school achievement and are based on the cumulative inputs before school. Because of difficulties in measuring achievement at the very beginning of schooling, however, entering achievement is taken to be achievement at the end of the first grade. The empirical specification does allow for differences among first-grade teachers, again employing a general covariance structure.

For preschool achievement, the models must be modified somewhat. Family inputs must be accumulated across time. This is done by

TABLE 4

INITIAL ACHIEVEMENT MODELS

Variable	Word Knowledge*	Reading Preparedness*
AVER KIDS*	−.097	−.055
	(−2.9)	(−1.7)
YOUNGER	.001	.004
	(1.0)	(.1)
NUMYNG*	−.071	−.112
	(−1.6)	(−2.7)
PERM*	.090	.091
	(2.8)	(3.0)
FEMALE	.060	.048
	(1.8)	(1.5)
Intercept	−.008	−.110
	(−.0)	(−.4)
TEACHERS	yes	yes
Degrees of freedom	230	230
R^2	.25	.20

* Indicates natural logarithm.

reconstructing the family composition through each child's preschool years and calculating the average number of children faced by each, using the specific family history on number of children and the spacing between each child. Furthermore, for preschoolers the optimization model is altered because the time constraint for dealing with children in and out of school is undoubtedly different from that in the simple theoretical derivation. If not working, mothers can spend time during the school day with preschoolers but not with schoolers, and thus the effective time constraint differs across children in the family. While all children can be viewed as competing for nonschool time, only nonschoolers compete for time during the school day. Therefore, consideration must be given to the age distribution of the children, and the existence and number of younger siblings are included in the estimated models.[9] Performance is measured by tests of word knowledge and reading preparedness—tests designed for younger students.

The effects of family inputs on preschool achievement, displayed in table 4, are very similar to those for achievement growth during

[9] Such considerations are clearly relevant in analyzing time allocations to both school and preschool children because of the possibility of sequencing the provision of private time to match each child's available time during the day. This is tested in the empirical work, and it appears that the school-preschool distinction matters only for preschoolers, i.e., that school time is a "bonus" to preschoolers and that sequencing of attention does not have a strong influence on the achievement of older children. The results on later grades differ from those of Stafford (1987), who finds that younger siblings, particularly boys, detract from performance. Further, in none of the estimates here does the sex composition of siblings matter.

the school years. Larger families depress achievement, and this effect is even stronger if there are younger siblings competing for attention. Higher permanent income leads to higher initial achievement.

The magnitudes of the estimated family effect elasticities are larger than those previously obtained in the achievement growth models. This is consistent with the longer time period for the effects to take place during preschool years as compared to individual school years.

As found in the previous models, work behavior of mothers or single-parent families have no effect on performance (although the number of working mothers is low). Further, child spacing per se has no effect on preschool achievement.

IV. Family Size, Birth Order, and Child Spacing

The previous results provide considerable confirmation of the basic dynamic time allocation model. Achievement growth through the schooling period is primarily affected by the number of "competing" siblings and the quality of parental time.

In the estimated value-added models, position in the birth order (being first born or last born), sibling spacing, and age structure of the family have no effect on achievement growth after total family size is considered. But these factors have an impact, even given nondiscriminatory time allocations, through altering the child-specific patterns of inputs.

The interplay of these factors becomes apparent when expected achievement differences are calculated for families of varying sizes and child spacing. The estimated models in tables 2 and 4 provide the basis for calculating the advantages of being in families of differing size and composition. Children of differing birth orders obviously face a different time pattern of family compositions so that expected achievement will vary systematically within families. For each completed family size and a given child spacing, comparisons are made to the expected achievement of an only child, with all other relevant attributes of families and schools held constant.

Figures 1 and 2 display the relative sixth-grade performance of children of differing birth order by completed family size.[10] The dominant factor is clearly total number of children, which is not par-

[10] These results are based on an average spacing of 24 months. Increased child spacing acts, as one would expect, like a reduction in family size. Spacing does tend to differ both between children in a family and across families of differing sizes; the median spacing is larger for the last child in any completed family size than for earlier children, and the average spacing will be less in larger families than in smaller families (U.S. Bureau of the Census 1984). The patterns discussed here hold across child spacings of 18–42 months, and the quantitative differences are relatively small with different spacing.

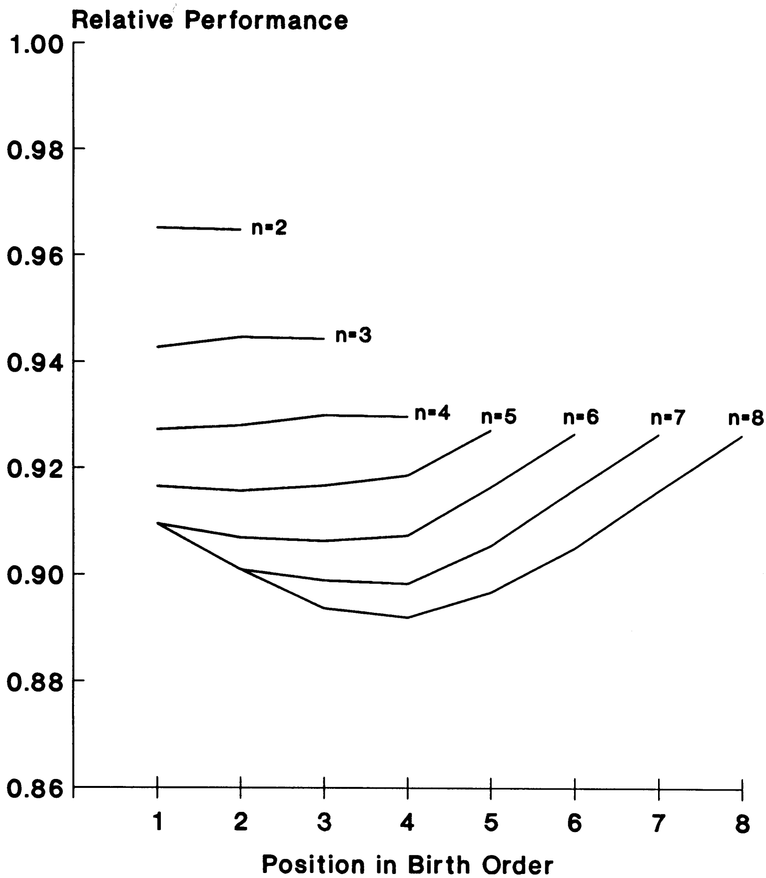

FIG. 1.—Sixth-grade reading performance relative to performance of only child by number of children in family (n) and by position in birth order.

ticularly surprising given the form of the final estimates. Within families—particularly smaller families—position in the birth order has a very minor effect on performance. Figure 1 indicates that, compared to an only child, the children in a two-child family would each achieve about 3.5 percent less in reading scores by the end of the sixth grade. The drop with increased family size is even more severe in vocabulary scores (fig. 2).

In large families, it becomes better to be last born rather than first born. The reason is simple: the first born has an advantage (a smaller family and thus more attention) early in life, but the last born has the same advantage later in life. Moreover, the last born does not have younger children competing for the preschool time of the mother. Indeed, as shown in the figures, the last born in families of four or

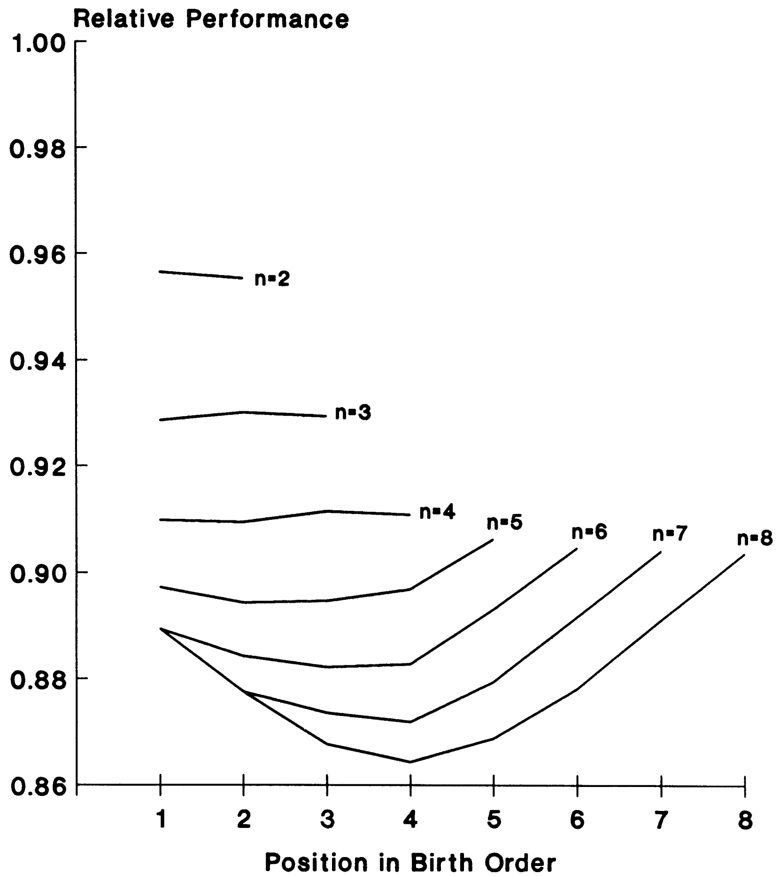

FIG. 2.—Sixth-grade vocabulary performance relative to performance of only child by number of children in family (n) and by position in birth order.

more children achieve at approximately the same level because they face approximately the same family configuration through their schooling years. Similarly, the next to last in families of five or more achieve at similar levels, regardless of completed family size, given the assumed child spacing. These achievement patterns are qualitatively similar to the time allocation patterns estimated by Lindert (1977).

Many previous authors have advanced reasons why first-born children outperform later-born children and, in fact, have produced evidence indicating higher performance by children earlier in the birth order. These explanations have involved special attention to first born, a variety of biological explanations, parental age, and so forth. The evidence here does not support these explanations.

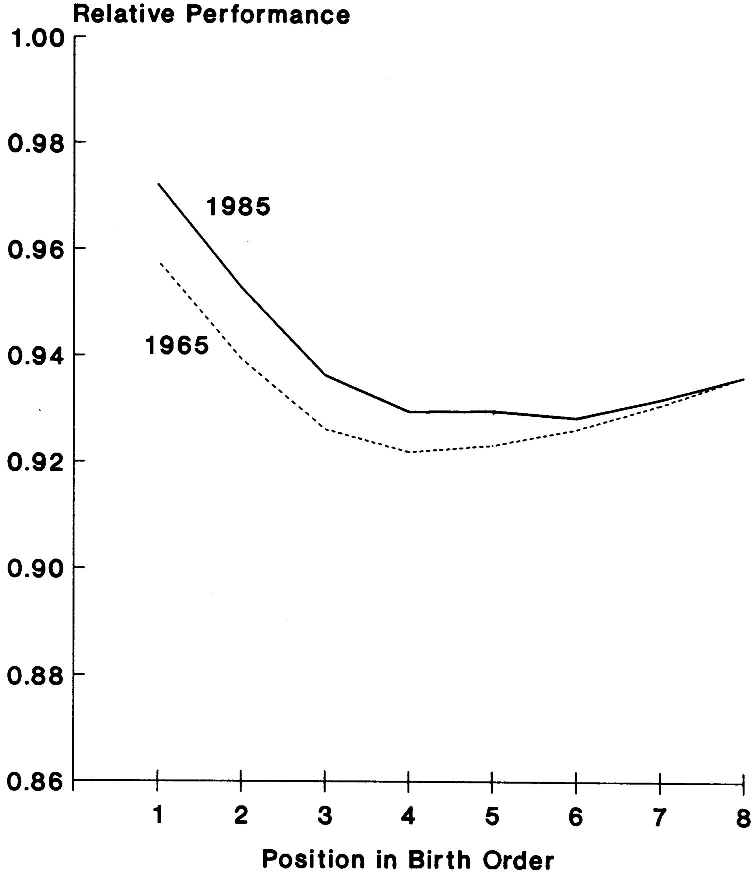

FIG. 3.—Sixth-grade reading performance relative to performance of only child by position in birth order alone: 1965 and 1985 family size distributions.

There is, nevertheless, a distinct advantage of being first born: The first born has a higher probability of being in a smaller family than those born later in the birth order. This advantage can be seen by merging data on the family size distribution with the performance data of figures 1 and 2. Figures 3 and 4 display the relative achievement of children at different places in the birth order based on the black family distributions in 1965 and 1985 (U.S. Bureau of the Census 1966, 1986). Being early in the birth order has a dramatic effect on performance, but this comes entirely from the family size effect.

Between 1965 and 1985, there was also a striking shift toward smaller families. For black families, the mean number of children for families having children fell from 2.93 to 1.99. This fall has the effect

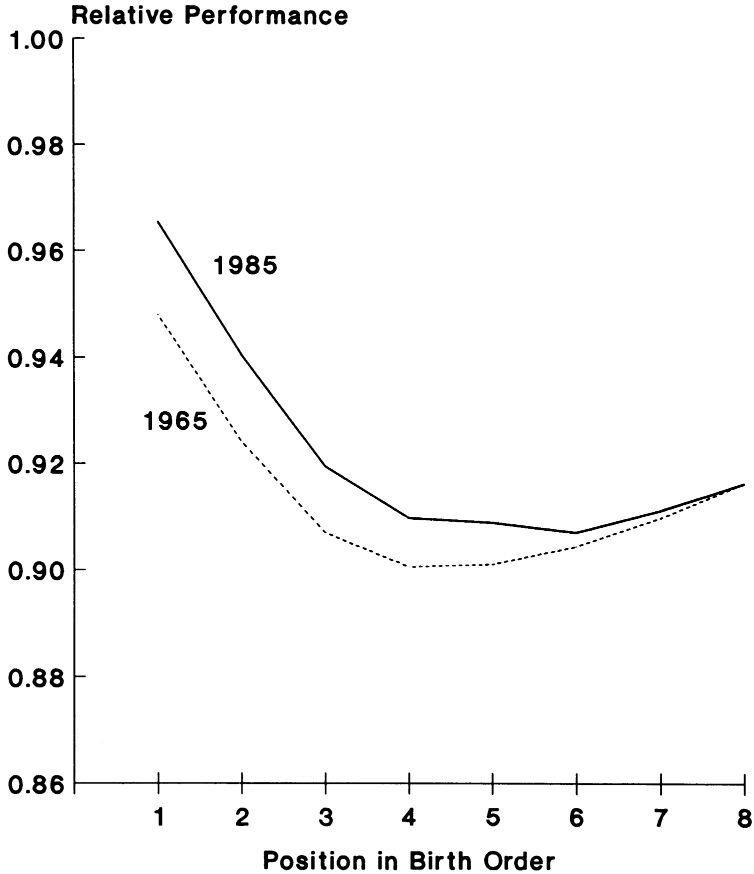

FIG. 4.—Sixth-grade vocabulary performance relative to performance of only child by position in birth order alone: 1965 and 1985 family size distributions.

of increasing the performance throughout the birth order and of flattening the relative performance curve. The potential magnitude of these changes is indicated by the comparisons of different family size distributions found in figures 3 and 4.

The sample used here—low-income black students from Gary, Indiana—is clearly not representative of the entire population. On the other hand, the available data are superior to those that have been available in virtually any other schooling situation. The longitudinal collection of detailed home and school information allows much more precise estimates of family composition effects than previously possible.

These data indicate that the aggregate effects of changing family composition probably have contributed to the observed aggregate changes in test scores, although the exact magnitude of any such effects is difficult to estimate precisely. The most noted change was the fall in the Scholastic Aptitude Test (SAT) that occurred between 1965 and 1980; however, this fall was observed in many other tests, including the Iowa achievement tests, over this period (Congressional Budget Office 1986). Some have suggested that this decline in academic performance was caused by the changes in family composition (e.g., Zajonc 1976; Easterlin 1978).

To illustrate the potential effects of a changing family size distribution, we can calculate the effect of the substantial sizable changes in family composition that occurred between the mid-1960s and the mid-1980s. (Note that these changes are virtually the mirror image of those affecting the SAT-taking population over the period and that the results overstate the effects on any specific birth cohort.)[11] The most straightforward interpretation of the results is that they indicate the potential effects of family size change for low-income, urban blacks. However, if we assume that the estimated models provide an index of the total effect that might be observed for a national sample, we can (with considerable caution) go further to bound the potential national effects.

Table 5 displays the aggregate achievement comparisons based on the alternative family size distributions of 1965 and 1985 by race. The steady-state movement from the large families of 1965 to the smaller ones of 1985 yields, by these estimates, a 1.5–1.9 percent improvement in achievement for the population as a whole. Blacks have had larger families, which are detrimental to their achievement, but their improvement over time is greater. Thus the achievement gap due to larger families has lessened noticeably.

The independent effects of family size and composition on achievement extrapolated in table 5 from this analysis, when converted to standard deviations of test performance, range from .063 to .085 standard deviations of *within-sample* sixth-grade Iowa reading and vocabulary test scores.[12] (This implies that the median person in the

[11] The calculations are best thought of as indicating the magnitude of achievement change when one goes between two steady states. The distribution of family sizes, which reflect the number of resident children aged 18 or under, in a given year includes children of 18 separate birth cohorts. When the distribution of families is changing over time, each of those birth cohorts will face a different pattern of family sizes, and the pattern existing at the time a cohort takes the SATs will not necessarily be a good indicator of the history faced by that cohort. Average family size rose through the 1950s until 1965 and subsequently fell. The cohort reaching 18 at the peak in family sizes will have experienced smaller family sizes over its schooling years than a later cohort that tends to be born into larger families and to experience these larger families throughout the school years.

[12] To compare across tests, it is convenient to translate the estimated effects into standard deviations. However, the within-sample standard deviations, which come

TABLE 5

AGGREGATE ACHIEVEMENT FOR ALTERNATE FAMILY SIZE DISTRIBUTIONS
(Relative to 1965 Distribution for U.S. Population)

Family Size Distribution	Vocabulary	Reading
1965 total U.S. population	1.000	1.000
1965 black population	.991	.988
1985 total U.S. population	1.015	1.019
1985 black population	1.011	1.011

NOTE.—Calculations based on 30-month child spacing.

1965 distribution would perform at the level of the forty-sixth to forty-eighth percentile in the 1985 distribution.) The decline in SAT scores from peak to trough for verbal was .48 standard deviations and for math was .28 standard deviations (Congressional Budget Office 1986). On the other hand, the decline in sixth-grade Iowa vocabulary tests for the nation as a whole was only .10 standard deviations and that for reading was .17 standard deviations.[13] While the use of within-sample standard deviations undoubtedly exaggerates the effects, family factors can clearly play a significant role in the aggregate test score changes. Changes in family size of the magnitude observed over the past two decades could explain half or more of the aggregate changes in sixth-grade Iowa achievement tests and 15–20 percent of the SAT verbal performance if the differentials remained through all grades.

V. Teachers, Schools, and Parental Choices

The bulk of this analysis has focused on the effects of family environment on student achievement. Nevertheless, while family inputs to education are indeed extremely important, the differential impacts of schools and teachers receive more attention when viewed from a policy viewpoint. This reflects simply that the characteristics of schools are generally more easily manipulated than what goes on in the family.

Differences among teachers are unquestionably large and significant, indicating their potential for decisively altering student achievement. Incorporating teacher dummy variables into the basic models

from the severely truncated samples in this analysis, will undoubtedly underestimate the population standard deviations.

[13] As described in Congressional Budget Office (1986), virtually all tests exhibited a decline in the late 1960s and into the 1970s, with the SAT changes being largest. The vocabulary test showed the most stability of all the Iowa tests, and the sixth-grade drop was less than that at higher grades.

for grades 2–6 (table 2) increases R^2 by a third. More important, *the estimated difference in annual achievement growth between having a good and having a bad teacher can be more than one grade-level equivalent in test performance.*[14] Students who have a "run" of good or bad teachers can thus end up in very different achievement positions after just a few years of schooling.

Several critical issues about differences among teachers are uniquely addressed with these data. First, the extended panel permits an initial investigation of the nature of estimated teacher differences. Second, when information is added about the characteristics of the teachers, the relationship between common policy approaches and student performance can be assessed. Finally, some preliminary consideration of the choices and actions of parents in selecting schools is possible.

A. *Teacher Skill Differences*

The covariance analysis yields direct estimates of the variation in mean achievement growth of students across different teachers and classrooms. All previous covariance analyses have observed teachers at only a single point in time, making it difficult to distinguish between the effect of the teacher and some possible interaction of the teacher and class. Nevertheless, while based on indirect evidence, the previous results have been interpreted as primarily reflecting differences in teacher "skill," that is, differences in the overall ability of teachers in eliciting learning growth of students (see Hanushek 1986). This interpretation comes from two sources. First, all the previous covariance estimates have employed value-added models. The estimated teacher effects have not been related systematically to the initial or entering achievement levels (A_{it*} in eq. [2]). This implies that any direct placement effect, such as putting the "best" teacher in the "best" or "worst" classes, would be extremely hard to implement. Second, principals' evaluations of teachers are highly correlated with the estimated teacher effects (see Murnane 1975; Armor et al. 1976). Both suggest that the estimated effects reflect differences among teachers and not the students that they have.

This sample permits a novel investigation of skill differences because teachers are observed at different points of time with different students. The strongest version of the skill hypothesis would be that

[14] A movement from one standard deviation below the mean to one standard deviation above the mean of estimated teacher effects is equal to approximately one standard deviation of student performance on both the reading and vocabulary tests. One standard deviation on the student tests ranges from one to one and a half grade-level equivalents, depending on the grade (see App. table A1).

teachers perform consistently across time, classrooms, and school settings. A weaker version is simply that there is a tendency for similar performance, but that certain specific factors can alter teacher performance. A distinct alternative is that teacher performance is highly related to the specific classroom and set of students that a teacher is given. This latter notion forms the basis for teachers' concerns about classroom assignment, particularly when faced with the possibility of merit pay. The disparate policy implications related to these alternative hypotheses imply that evidence on them is very valuable.

Teachers elicit similar growth across the two areas of achievement, reading and vocabulary. The hypothesis that teacher effects are the same across the two output measures, tested by pooling the samples across tests, is only barely rejected; the F-statistic for equality equals 1.42, and the critical value for the 5 percent level is 1.20.[15] The similarity across measured outputs is consistent with a skill interpretation but does not distinguish clearly among the alternative interpretations.

The more interesting investigations involve the stability of teacher effects over time. The sampling design, which considers student achievement over four different academic years, implies that some teachers appear repeatedly in the different years with different students and, thus, that the source of teacher and classroom differences can be investigated further. When the teachers reappear in different years, they may be teaching at the same or different grade levels. While the sample is stretched considerably for this analysis, there are 22 teachers who are observed in multiple grades and have at least three students in each grade. (One of these teachers is found in three separate grades, and another is found in four separate grades.) An F-test for equality of individual teacher effects across different grades leads to rejection of the hypothesis for both vocabulary and reading achievement, but this rejection appears to come from a small number of teachers who exhibit large performance differences across grades. (The F-statistics are 3.62 and 4.26 for vocabulary and reading, respectively; the critical value for the 5 percent level is 1.52.) Altogether, five out of the 22 teachers were found to have statistically significant differences across grades on each test. Only two teachers, however, had significantly different impacts in both reading and vocabulary performance, indicating something obviously different for these teachers as opposed to simply a sampling phenomenon.

An additional 39 teachers are observed in different years but in the same grade. While equality of teacher effects across years is also

[15] The samples for the two output measures are pooled; the teacher coefficients are constrained to be the same, and the other exogenous factors are allowed to vary across tests. No allowance was made for the fact that individual students appeared twice in the pooled sample. The relevant degrees of freedom for the test are 182 and 3,834.

rejected, the differences by year appear less important than the grade-level differences. (The F-statistics for equality of all teacher coefficients are 3.43 and 2.41 for vocabulary and reading, with a critical value of 1.35 at the 5 percent level.) In terms of individual teachers, 12 were significantly different across years in vocabulary growth and eight in reading growth. But, again, only four teachers had significantly different classroom effects for both tests.

This work represents the first direct investigation designed to separate teacher skill effects from more general classroom differences. While there are some variations in teacher performance across grades and across years, there is an underlying stability in the estimated effects, suggesting support at least for the weaker version of the skill difference hypothesis. The samples do, nonetheless, become quite thin when such stability issues are investigated, and it is not possible to pursue the causes behind the few clear cases of distinctly different performance.

B. Teacher Characteristics

The more traditional approach to investigating school and teacher differences involves basing estimation on a small number of measured attributes such as class size or teacher education level. The data collection scheme for the Gary sample obtained survey information from the elementary school teachers of sampled students. These data permit estimation of traditional models, although the student sample is cut approximately in half by nonresponse to the teacher questionnaire. A comparison of the teacher skill estimates from the covariance analysis for the entire sample indicates no mean difference in skill for those responding and not responding to the teacher survey. The t-statistics for equality of mean teacher effects across response categories are 1.02 and 0.59 for the vocabulary and reading tests, respectively.

Table 6 displays estimated effects of teacher attributes that are consistent with previous analyses. Of the determinants of teacher expenditures per pupil (i.e., teacher experience and degree level and class size), only years of teacher experience are significantly related to student performance.[16] Previous estimates (reviewed in Hanushek [1981, 1986, 1989]) arrive at similar conclusions, suggesting serious questions about any relationship between expenditure variations and student performance.

[16] Teacher salaries are determined by the level of teachers' education and experience. Thus, combined with class size, these three factors determine variations in instructional expenditures per student.

TABLE 6

RELATIONSHIP OF SPECIFIC TEACHER CHARACTERISTICS AND ACHIEVEMENT GROWTH

Variable	Vocabulary*	Reading*
KIDS*	−.032	−.022
	(−1.8)	(−1.3)
PERM*	.008	.016
	(.5)	(1.1)
FEMALE	−.010	.027
	(−.6)	(1.8)
FAMTEST	−.012	−.014
	(−.6)	(−.7)
RELFAM*	.105	.038
	(2.4)	(1.0)
FALL	−.272	−.215
	(−7.1)	(−5.9)
	Teacher Attributes	
EXPER*	.070	.049
	(5.2)	(3.9)
MASTER'S	.005	−.001
	(.3)	(−.0)
CLASS SIZE*	.070	.014
	(1.4)	(.3)
FEM TEACHER	.024	.031
	(.8)	(1.1)
WHITE	−.076	−.071
	(−2.3)	(−2.3)
TEST*	−.001	.186
	(−.0)	(3.1)
Intercept	.424	−.417
	(1.1)	(−1.1)
Degrees of freedom	996	996
R^2	.581	.614

NOTE.—Models include dummy variables for grade levels and log of pretest scores (not shown). Variables not defined in table 1 are the following: EXPER: years of total teacher experience; MASTER'S: dummy variable that equals one if teacher has master's degree, zero otherwise; CLASS SIZE: number of students on class register; FEM TEACHER: dummy variable that equals one if teacher is female, zero otherwise; WHITE: dummy variable that equals one if teacher is white, zero otherwise; TEST: teacher's score on *Quick Word Test.*

* Denotes logarithms of variables.

Within this sample, there is no indication of differences in performance for male and female teachers, but white teachers do significantly worse than black teachers. The sampled students are all black. Therefore, this result may reflect either that black students do better with teachers of their own race or that the white teachers that are attracted to this setting are otherwise poorer, given their measured characteristics. Finally, teachers took a short word test, which is frequently interpreted as a substitute for a general IQ test. There is mixed evidence about the relationship of teachers who score higher and student performance. "Smarter" teachers appear to do better in improving reading performance, but not vocabulary performance.

This ambiguous result is, however, similar to previous studies that find inconsistency in the relationship with teacher tests (Hanushek 1986).

As suggested in previous analyses (Hanushek 1986, 1989), policies aimed at altering simple organizational and hiring practices are unlikely to be effective. Reducing class sizes, which has been a policy pursued vigorously for several decades, has no apparent effect on student performance. Similarly, requiring teachers to have a master's degree—as many states do—just increases costs, not student performance. The large differences in teacher performance described in the previous section clearly are not captured by simple descriptors of teachers and classrooms. The factors that are found to be significant in this study (teacher experience and, for reading, teacher test scores) have not been consistently important across studies and, moreover, capture only a limited portion of the total variation in teacher performance estimated here. The implication of the analysis is that entirely different incentive structures for schools and teachers must be considered (see, e.g., Chubb and Hanushek 1990), but such issues are outside of the scope of this paper.

C. *Parental Choices*

The influence of families is not restricted to just the amount of education in the home but also includes parents' interactions with schools and the selection of teachers. It is frequently argued that parents, particularly middle- and upper-income parents, search intensively for good schools. Further, search behavior of parents is particularly important in consideration of educational vouchers and other market mechanisms for schools. Nevertheless, very little is actually known about this search process.

Here I present an exploratory analysis into how variations in teacher quality are related to family characteristics. The specific issues considered include whether higher-income families, families with two parents, families with working mothers, and families for which there is a change of schools systematically find good teachers for their children. The quality of each student's teacher, as measured by the mean student achievement growth associated with each teacher, is regressed on the various family measures. While the previous analyses considered how, say, having a working mother might affect the individual student's achievement (with family background and teacher quality held constant), this analysis considers whether or not the children of working mothers systematically get placed in classes with the better teachers.

The results, presented in table 7, do not provide any indication of

TABLE 7

PARENTAL CHOICE OF TEACHERS

Variable	Vocabulary	Reading
MOTHER WORKS	.0058	.0067
	(.6)	(.9)
MALE PRESENT	−.0084	−.0026
	(−.9)	(−.3)
SCHOOL CHANGE	−.0209	−.0055
	(−1.3)	(−.4)
PERM ($1,000's)	.0001	.0001
	(.0)	(1.1)
FALL	−.0103	.0011
	(−3.2)	(.3)
Intercept	−.0210	.7817
	(−.1)	(3.3)
R^2	.007	.002

NOTE.—Dependent variable is estimated teacher quality from models in cols. 1 and 4 of table 2. SCHOOL CHANGE equals one if student changed schools and zero otherwise.

systemic relationships with these factors. The only variable that is statistically significant is the measure of whether the vocabulary pretest was taken in the fall (FALL), which is included only to correct teacher quality estimates for potential biases from the different testing time. Again, however, it must be remembered that this is a sample of low-income households. While variations in income within this restricted range are not related to teacher choice, this might not hold over a broader range.

These results suggest that investigations of teacher choice will require much more detailed information about family motivations and behavior. Further, since schools in this sample did not provide for active choice, these results also include the actions of schools, and those actions might well mask any choice behavior by parents.

VI. Conclusions

A distinct trade-off between quantity and quality of children is found to exist. The theoretical model, extending the basic analyses in economic demography, considers the allocation of time to children and describes the implications of alternative within-family allocation schemes. The empirical analysis finds that achievement falls systematically with increased family size. Further, parents appear to act in a compensatory manner, favoring lower-ability children within the family, or in a neutral manner. There is no evidence of achievement maximization by parents, although the sample is restricted to low-income black families and students. The results are consistent with a maximizing model of parental time allocation. Since actual allocations are not observed, however, other interpretations might be possible.

Families differ significantly in terms of the quality of inputs as measured by permanent income. There is, at the same time, no evidence that changing the immediate circumstances of the family will have any effect on student performance. The work behavior of the mother has no influence on the educational performance of children. Neither does the absence of a father. These findings, which are generally consistent with other research (Congressional Budget Office 1987), are encouraging since they indicate that the massive societal changes of the past two decades will not have a noticeable negative effect on human capital formation.

While it is always better to be in a smaller family, there is no particular advantage to being first or last born when family size is held constant. However, necessarily, being first born increases the chances of being in a small family. Therefore, the average first born will outperform the average second born, and so forth through the birth order. The dramatic changes in family size and composition of the past two decades are shown to be large enough to affect aggregate performance noticeably. In terms of the Iowa tests, the movements in family size could potentially explain over half of the observed (peak to trough) change.

The analysis of teacher effects on student performance provides a number of new results. Specifically, there is no doubt that teachers vary dramatically in effectiveness. The *difference* in student performance in a single academic year from having a good as opposed to a bad teacher can be more than one full year of standardized achievement. There is additional support for a teacher skill interpretation of differences in classroom performance of students. This support comes from the general stability of teacher impacts over time and grades. Additionally, measured characteristics of teachers do not capture skill differences very well, corroborating previous analyses of the overall inconsistency of results for individual teacher characteristics. Finally, an exploratory analysis of parental choices of teachers does not yield any systematic behavioral outcomes.

Appendix

TABLE A1

MEANS AND STANDARD DEVIATIONS (in Parentheses) BY GRADE LEVEL

Variable	1	2	3	4	5	6
READING:						
POST	1.92	2.51	3.31	3.99	4.43	5.11
	(.6)	(.7)	(1.0)	(1.1)	(1.2)	(1.2)
PRE	. . .	1.98	2.54	3.26	3.86	4.71
		(.6)	(.7)	(1.0)	(1.1)	(1.3)
VOCABULARY:						
POST	2.04	2.97	3.43	4.17	4.65	5.45
	(.6)	(1.1)	(1.0)	(1.2)	(1.5)	(1.6)
PRE	. . .	2.13	2.93	3.39	3.96	4.76
		(.7)	(1.0)	(.9)	(1.2)	(1.5)
KIDS	4.29	4.41	4.47	4.47	4.46	4.61
	(2.3)	(2.3)	(2.1)	(2.1)	(2.1)	(2.2)
MOTHER'S EDUCATION (years)	10.61	10.55	10.62	10.57	10.40	10.18
	(1.6)	(1.7)	(1.7)	(1.7)	(1.7)	(1.8)
PERM	$5,877	$6,135	$6,628	$6,492	$6,499	$6,619
	(2,674)	(2,777)	(3,107)	(3,059)	(3,041)	(3,013)

MOTHER'S HOURS:						
All mothers	3.87	4.07	5.22	3.87	4.62	5.69
	(10.4)	(10.7)	(12.4)	(10.8)	(11.6)	(12.9)
Mothers with job	24.01	26.04	28.72	27.73	28.09	30.11
	(13.7)	(12.6)	(13.3)	(13.5)	(12.9)	(12.0)
CHILD'S AGE	7.24	8.32	9.39	10.46	11.53	12.59
	(.4)	(.5)	(.5)	(.5)	(.6)	(.6)
MOTHER WORKS	.16	.16	.18	.14	.16	.19
MALE PRESENT	.34	.34	.42	.39	.39	.40
MALE CHANGE	.08	.08	.12	.06	.06	.04
SCHOOL CHANGE	.05	.06	.05	.04	.04	.05
School year:						
1972	...	...	...	...	...	.27
1973	.59	...	.36	.37	.35	.19
1974	.34	.75	.39	.36	.39	.31
1975	.07	.25	.25	.26	.26	.24
Reliability coefficients:*						
Reading	.92	.92	.93	.93	.93	.91
Vocabulary	.87	.88	.88	.89	.89	.90
Number of observations	210	192	391	415	441	473
Number of teachers	95	107	114	113	101	84
With 5 or more kids	10	12	22	26	32	31
With 3 or more kids	28	25	41	46	44	47

* Source: Iowa Tests of Basic Skills, *Manual for Administrators (Grades 3–6)*, 1964.

References

Armor, David, et al. *Analysis of the School Preferred Reading Program in Selected Los Angeles Minority Schools.* Report no. R-2007-LAUSD. Santa Monica, Calif.: Rand Corp., 1976.

Becker, Gary S. "An Economic Analysis of Fertility." In *Demographic and Economic Change in Developed Countries.* Universities–National Bureau Conference Series, no. 11. Princeton, N.J.: Princeton Univ. Press (for NBER), 1960.

———. *A Treatise on the Family.* Enl. ed. Cambridge, Mass.: Harvard Univ. Press, 1991.

Becker, Gary S., and Lewis, H. Gregg. "On the Interaction between the Quantity and Quality of Children." *J.P.E.* 81, no. 2, pt. 2 (March/April 1973): S279–S288.

Becker, Gary S., and Tomes, Nigel. "Child Endowments and the Quantity and Quality of Children." *J.P.E.* 84, no. 4, pt. 2 (August 1976): S143–S162.

———. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *J.P.E.* 87 (December 1979): 1153–89.

Behrman, Jere R.; Pollak, Robert A.; and Taubman, Paul. "Parental Preferences and Provision for Progeny." *J.P.E* 90 (February 1982): 52–73.

———. "Do Parents Favor Boys?" *Internat. Econ. Rev.* 27 (February 1986): 33–54.

———. "Family Resources, Family Size, and Access to Financing for College Education." *J.P.E.* 97 (April 1989): 398–419.

Behrman, Jere R., and Taubman, Paul. "Birth Order, Schooling, and Earnings." *J. Labor Econ.* 4, no. 3, pt. 2 (July 1986): S121–S145.

Blake, Judith. *Family Size and Achievement.* Berkeley: Univ. California Press, 1989.

Chubb, John E., and Hanushek, Eric A. "Reforming Educational Reform." In *Setting National Priorities: Policy for the Nineties,* edited by Henry J. Aaron. Washington: Brookings Inst., 1990.

Congressional Budget Office. *Trends in Educational Achievement.* Washington: Congressional Budget Office, 1986.

———. *Educational Achievement: Explanations and Implications of Recent Trends.* Washington: Congressional Budget Office, 1987.

Cronbach, Lee J., and Furby, Lita. "How Should We Measure 'Change'—or Should We?" *Psychological Bull.* 74 (July 1970): 68–80.

Easterlin, Richard A. "What Will 1984 Be Like? Socioeconomic Implications of Recent Twists in Age Structure." *Demography* 15 (November 1978): 397–432.

Hanushek, Eric A. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions." *J. Human Resources* 14 (Summer 1979): 351–88.

———. "Throwing Money at Schools." *J. Policy Analysis and Management* 1 (Fall 1981): 19–41.

———. "The Economics of Schooling: Production and Efficiency in Public Schools." *J. Econ. Literature* 24 (September 1986): 1141–77.

———. "The Impact of Differential Expenditures on School Performance." *Educ. Researcher* 18 (May 1989): 45–51.

Harbison, Ralph W., and Hanushek, Eric A. *Educational Performance of the Poor: Lessons from Rural Northeast Brazil.* New York: Oxford Univ. Press (for World Bank), 1992 (in press).

Hayes, Cheryl D., and Kamerman, Sheila B., eds. *Children of Working Parents: Experiences and Outcomes.* Washington: Nat. Acad. Press, 1983.

Hayes, Donald P., and Grether, Judith. "The School Year and Vacations: When Do Students Learn?" Manuscript. Washington: Urban Inst., 1969.

Heyns, Barbara L. *Summer Learning and the Effects of Schooling.* New York: Academic Press, 1978.

Hill, C. Russell, and Stafford, Frank P. "Allocation of Time to Preschool Children and Educational Opportunity." *J. Human Resources* 9 (Summer 1974): 323–41.

———. "Parental Care of Children: Time Diary Estimates of Quantity, Predictability, and Variety." *J. Human Resources* 15 (Spring 1980): 219–39.

Kamerman, Sheila B., and Hayes, Cheryl D., eds. *Families That Work: Children in a Changing World.* Washington: Nat. Acad. Press, 1982.

Leibowitz, Arleen. "Home Investments in Children." *J.P.E.* 82, no. 2, pt. 2 (March/April 1974): S111–S131.

Lindert, Peter H. "Sibling Position and Achievement." *J. Human Resources* 12 (Spring 1977): 198–219.

Murnane, Richard J. *The Impact of School Resources on the Learning of Inner City Children.* Cambridge, Mass.: Ballinger, 1975.

Murnane, Richard J.; Maynard, Rebecca A.; and Ohls, James C. "Home Resources and Children's Achievement." *Rev. Econ. and Statis.* 63 (August 1981): 369–77.

Ragosa, David R., and Willett, John B. "Understanding Correlates of Change by Modeling Individual Differences in Growth." *Psychometrika* 50 (June 1985): 203–28.

Stafford, Frank P. "Women's Work, Sibling Competition, and Children's School Performance." *A.E.R.* 77 (December 1987): 972–80.

U.S. Bureau of the Census. "Household and Family Characteristics: March 1965." *Current Population Reports,* ser. P-20, no. 153. Washington: Government Printing Office, August 5, 1966.

———. "Childspacing among Birth Cohorts of American Women: 1905 to 1959." *Current Population Reports,* ser. P-20, no. 385. Washington: Government Printing Office, February 1984.

———. "Household and Family Characteristics: March 1985." *Current Population Reports,* ser. P-20, no. 411. Washington: Government Printing Office, September 1986.

Willis, Robert J. "A New Approach to the Economic Theory of Fertility Behavior." *J.P.E.* 81, no. 2, pt. 2 (March/April 1973): S14–S64.

Zajonc, R. B. "Family Configuration and Intelligence." *Science* 192 (April 16, 1976): 227–36.

Zajonc, R. B., and Markus, Gregory B. "Birth Order and Intellectual Development." *Psychological Rev.* 82 (January 1975): 74–88.