

Chapter 9:

Satisfaction with Schools

From the market standpoint, the foundation upon which so much of the argument for school choice rests, increased consumer satisfaction is an expected outcome. Simply put, people who have choices should choose things they think best meet their needs and, in turn, they should be more satisfied with them. By extension, parents who choose their children's schools should be more satisfied than parents who do not.

In this chapter, we begin our empirical investigation of the effects of charter schools based on the survey data described in Chapter 3. Here we look at the extent to which parents and students are satisfied with their schools. While we have panel data—repeated measures on the same survey respondents over time—we begin our investigation with cross-sectional analyses.¹ We devote a great deal of space to the question of satisfaction for two reasons.

First, the structure of our data is complicated and we use several methods to account for potential problems for inference common to most observational studies, such as self-selection to the policy “treatment” and missing data due to nonresponse.

Our other reason for devoting so much time to the issue of how parents view their child's school has to do with the importance of parental satisfaction to theories of choice. While much of the current debate about school choice focuses on achievement gains, as noted above, greater parental satisfaction with schools of choice has been a long-standing

¹ Fundamental to our panel design is the idea that the effects of choice are dynamic and that therefore time matters. When we start exploring the effects of charter schools over time, we need to explore alternative methods for addressing dynamic effects, especially given the difficult nature of our panel—and we put the discussion of those problems and models off until the next chapter.

outcome that advocates have used to buttress their argument for expanding choice. If parents in charter schools do not prefer them to the traditional public alternative, this may signify that differences between the two sectors are illusory, or at best, that there is a serious shortfall in the range and type of charter schools available.

Satisfaction is also critical to the politics of school reform. Many charter school advocates have no qualms in talking about a charter school *movement*, linking charter schools to a larger political effort to reform the system of education in the United States. If this movement is to have traction and gain adherents, parents who are satisfied with their children's schools represent a potential pool of supporters (and, not incidentally, voters) for further expansion of charter schools and choice in general. In contrast, if charter school parents are not satisfied with their schools, then an important foundation for the "movement" is missing.²

Measuring How Parents Evaluate Their Child's School

In the first of this pair of chapters exploring satisfaction, we present some basic information on the relationship between the sector in which a child is enrolled and how parents evaluate that child's school. As noted earlier, in this chapter we confine ourselves to cross-sectional analysis. We use a broad set of measures of parental evaluation of their children's schools and we show a consistent pattern: Controlling for many factors that may affect the relationship between choice and satisfaction, charter school parents consistently evaluate their children's schools more highly than do parents with children in the traditional DC public schools.

² This argument is central to Moe's analysis of vouchers (Moe 2001), also see Clarence Stone (1998) on the importance of building civic capacity to sustain school reform.

Many of the factors that we control in our empirical analysis are the “usual suspects”—socioeconomic status, education levels, church attendance, and so on. However, one of the factors that must be controlled may be less evident but is embedded in the very nature of charter schools as schools of choice. As noted earlier in this book, charter schools are an “option demand” form of choice, in which parents are free, but not required to choose. In this two step process, parents first choose to choose and only then do they choose a school. Given the nature of this process, the parents who choose to choose are “self-selecting” into the charter school system and may not be representative of the entire population of parents. In turn, the characteristics that are motivating them to choose may affect their subsequent behavior and attitudes toward the schools.

To the extent this is true, a self-selection bias may enter into the data and simple comparisons of choosers and non-choosers will be affected by these parental characteristics independent of the quality of the charter schools themselves. That is, higher evaluations of charter schools may be a function of the factors that led charter school parents to choose in the first place. This problem is increasingly well-known to social science researchers and a variety of techniques have been introduced to deal with this potential self-selection bias.

While our substantive task in this chapter is to explore differences in satisfaction with charter schools, an equally important task is methodological: to see how well any observed differences stand up to controls for self-selection into charter schools and to tests of sensitivity for hidden bias due to unmeasured covariates that may be predictors of both opting to choose and satisfaction. In this chapter, we also examine the possibility that any observed positive evaluations given to charter schools may be the result of a

form of motivated reasoning (a topic which we briefly discuss in Chapter 8 in a different context). As we discuss in more detail below, the potential problem is simple to state: If a parent chooses her child's school, she may need to believe that the choice was the right one. In turn, she may view the chosen school through what Erikson calls "rose colored glasses." (Erikson 1982). Later in this chapter, we investigate the likelihood that the very act of choosing charter schools creates such a positivity bias. While social scientists have developed a wide range of tools to fix the self-selection problem, the potential bias from donning rose-colored glasses is less well-studied.

Choice and Charter Schools Have Been Linked to Higher Satisfaction

In our introductory chapters, we documented the rapid increase in the number of charter schools and the number of students enrolled in them. One likely reason for this growth is that parents and students think charter schools are better than the traditional public schools in which they were previously enrolled. For example, Finn et al. (1997) find that a large majority of parents feel that charter schools in which their children were enrolled are better than the traditional public schools they left, with respect to class size, school size, teacher attentiveness, and the quality of instruction and curriculum. In contrast, less than five percent of these parents found charter schools inferior. Finn et al. also find high levels of student satisfaction across a gamut of school attributes, including teachers, technology, class size, and curriculum. Teachers also seem to like charter schools, with high levels of satisfaction found among charter school teachers (Koppich 1998).

There are several possible foundations for this greater satisfaction with charter schools. Perhaps the strongest is that of "allocative efficiency"— education is a complex,

multifaceted “good,” and choice allows parents to select schools that emphasize the kind of education they want for their children (Schneider et al. 2000).³ This foundation for higher parent satisfaction in schools of choice dates at least as far back as Milton Friedman’s original argument in favor of vouchers in the 1950s and has been used to support choice ever since.

In his pioneering work, Friedman made a strong case for consumer sovereignty, arguing that higher levels of satisfaction with schools will flow from maximizing the freedom of parents to choose schools (Friedman 1955). From this perspective, choice leads to higher parental evaluation of choice schools because it increases the ability of parents to match their preferences for specific values, needs or pedagogical approaches with the school. As Goldring and Shapira put it: “The family sovereignty position suggests choice leads to greater satisfaction in that it accommodates individual family preferences, mainly in the areas of curricula, teaching philosophy, and religion. Parents will be satisfied in exercising their fundamental right of individual choice and freedom of belief about the best education for their children.” (1993, 397; see also Coons and Sugarman 1978 and Raywid 1989).

In addition to increasing this match between consumer and provider, choice may change the schools themselves—making better “products” available for parents to choose among. Indeed, fundamental to the push for choice is the idea that choice unleashes competitive pressure on the schools that makes them improve—and charter schools are often seen as a central tool to leverage such change (see, e.g., Hill 1997, Teske et al. 2000, Hoxby 2000c).

³ Many of the factors discussed here are associated with choice in general, and have not been developed specifically for charter schools as a form of choice—but clearly apply to the charter school option as well.

While the debate still rages about the effect of choice on academic outcomes, there are other outcomes from choice that are less contested—any of which can increase parental satisfaction. For example, many charter schools are designed to change the relationship between administrators, teachers, parents and students, to foster what Coleman (1988) refers to as “functioning communities.” In these communities, the tighter links from the school to parents, families, and students is associated with better educational experiences and all parties, including teachers, will be more satisfied (Driscoll 1993). This link underlies the basic findings developed in the research on “effective schools,” which shows that good interpersonal relations between members of the school community and shared beliefs and values combine to promote good teaching and a positive learning environment (see especially Bryk and Schneider 2002).⁴

Indeed, many charter schools have a culture (and sometimes even a written contract) that provides parents opportunities to influence school management and to become more involved with the processes of school governance and functioning (see, for example, Peterson and Campbell 2001; Finn et al. 1997). To the extent that this occurs, parental evaluations of their children’s schools should improve (Chubb and Moe 1990; Raywid 1989; Goldring and Shapira 1993; but see Benveniste et al. 2003, chapter 3, on the possibility that such contracts may function as ceilings on participation).

It is also important to remember that choice seeks not only to empower parents but also to change the role of students, making them more central in the design of education programs and in the functioning of the schools. In short, one goal of choice is to increase the attention paid by schools to student needs (Hill et al. 1997). In turn, choice

⁴ We explore these issues in more detail in Chapter 11.

seems to improve student-teacher relations. For example, in her study of the effects of choice, Driscoll found that choice students were more likely to report “they got along well with teachers, that the quality of teaching was high, and that teachers praised them and listened to them.” (1993, 158). In a similar vein, Finn et al. (1997) found that large numbers of charter school students liked the “good teachers” in their schools, who, according to these students, teach until they learn the material and who don’t let students fall behind. To the extent that this behavioral change strengthens the ties between students and teachers and increases the level of student satisfaction with the schools, parental satisfaction with the schools should in turn increase.

Finally, choice may put pressure on administrators, teachers and staff to be more “consumer friendly.” As Hassel observes: “charter schools cannot take their ‘customers’ for granted. Their very survival depends on the degree to which families believe the schools are responding to family preferences and working hard to provide the education they demand.” (Hassel 1999, 6; also see Teske et al. 2000 and Buckley and Schneider 2004). Thus, the competitive pressures on charter schools should increase their responsiveness to parent demands—and such responsiveness should lead to higher evaluations.⁵

In short, there is reason to believe that parents with children in charter schools should evaluate their schools more highly. In the first part of this chapter we explore this relationship among the parents in our study. We begin by looking at a way of evaluating

⁵ This notion that “consumers” of education will have more voice has long been a centerpiece of the market model of schooling (e.g. Chubb and Moe 1990). However, as Benveniste et al. (2003, chapter 3), echoing Henig (1994), point out public schools cannot necessarily take their clientele for granted, either, since parents as citizens have numerous rights in regards to their child’s school and recourse to the legal system if they are infringed. Indeed this debate about the relative efficacy of parent-citizens versus parent-consumers is at the very core of the theoretical arguments over school choice.

schools that should be familiar to readers: we asked parents to assign letter grades ranging from A to F to their child's school overall and for three specific aspects of the school: their child's teacher, principal, and school facilities.

A Caveat about Using Grades

Before proceeding with the analysis, note that there is a well-known pattern when using grades as measures of parent evaluations of schools—parents almost always give high grades to their children's schools. For example, Phi Delta Kappa regularly asks a sample of parents to grade their children's schools. In their 2003 survey, the magazine found that 48% of respondents assigned an A or a B to schools in their community with an additional 31% assigning the grade of C. This pattern of high grades has remained remarkably durable over time. Further, the grades parents give to their own children's school are consistently higher than the grades they assign to the nation's schools as a whole: The number of A's and B's rises to 55% for public school parents and to 68% for parents asked to grade the public school their oldest child attends (Gallup 2003).

However, we believe that this pattern is not important for our analysis, since we are interested in comparing parents in the charter schools and the traditional public schools in the same city, and ultimately we do control for conditions that could affect parents in the two sectors differently.

Charter Schools Receive Higher Grades

We start with a simple exploration of the distribution of these evaluative grades in our first wave of interviews. We then explore the extent to which the observational nature of our study may be confounding these results.

In Figure 9.1, we present the marginal distributions of grades using the data from the first wave of the survey. In the upper left corner, we graph the overall grades parents give to their child's school by enrollment in charter school or traditional DC public schools. We see that 49 percent of charter parents gave their child's teachers a grade of A, fully 10 percent more than parents whose children were in DCPS. At the opposite end of the spectrum, 3 percent of DCPS parents gave their child's school an F, while only 1 percent of charter school parents gave this failing grade. The differences displayed in Figure 9.1 are all significant at $p < .01$ level.⁶

Figure 9.1 here

We find virtually the same pattern for parental evaluations of teachers and principals (the upper right and lower left corners of Figure 9.1), where charter parents were again 10 percentage points more likely to give an A than DCPS parents and even more so for facilities, where 17 percent more charter parents gave a grade of A compared to DCPS parents. Thus on each of these three specific dimension of schooling, there is a clear shift in the distribution of grades in favor of charter schools.

We can we convert these letter grades into numerical ones (using the familiar A=4, B=3, ..., F=0 scale) and create an overall Grade Point Average ("GPA"). Using this summary score, charter schools overall earn a GPA of 3.3, about 10 percent higher (or a third of a "letter grade") than the 3.0 of DCPS schools.

⁶ The survey data are weighted for probability of inclusion and post-stratified on charter enrollment due to intentional oversampling of this subpopulation, as described in Chapter 3. All the figures, tables and models estimated below use the same weights.

Are These Differences a Function of Self-Selection Biases?

The results shown in Figure 9.1 do not control for any characteristics of the parents in our sample. As we have noted earlier, the option demand nature of charter schools may lead to a selection bias in who chooses: charter school parents are not likely to be a random sample of the wider population of parents. A simple way to test the robustness of the charter school effect on satisfaction is to estimate a linear regression model of the grades on likely confounding covariates. Our first models are independent linear regressions of the four grades on the following characteristics:

- a charter school indicator, coded 1 if the respondent's child is in a DC charter school;
- a set of three dichotomous variables for self-reported race (Hispanic, white, other, with African-American the excluded—and modal—category);
- residential mobility measured by the number of years the person has lived in her current neighborhood and by the years the person has lived in D.C.;
- whether or not the respondent was employed;
- whether or not the respondent reports that she is a frequent volunteer at her child's school;
- respondent's PTA membership;
- the respondent's marital status;
- frequency of church attendance (a seven category variable, treated here as continuous);
- the number of people that the respondent reports frequently discussing her child's education with;

- the number of non-school organizations that the respondent reports membership in;
- the grade the respondent assigned to the DC public schools in general;
- respondent's years of schooling and years of schooling squared;
- how long the respondent's child has attended the school, and;
- the child's grade level.

While most of these covariates are included for obvious reasons, several require further explanation. A particular issue in evaluating a policy program such as charter schools with observational data is that participants in the program may differ in important and difficult-to-measure psychological dimensions such as motivation or dissatisfaction with previous choices. In this first approach, we attempt to control for this possibility with their school participation (volunteerism, PTA membership) and overall grade they assign the DC schools (general dissatisfaction). We return to this issue of controlling for motivation below, but first we present our initial model.

Let Y_{ig} denote respondent i 's grade assigned to each dimension of schools we measured (g). We thus estimate four independent models of the form:

$$\begin{aligned}
 Y_{ig} = & b_0 + b_1 \text{Charter}_i + b_2 \text{White}_i + b_3 \text{Hispanic}_i + b_4 \text{Other Race}_i \\
 & + b_5 \text{Years Lived in D.C.} + b_6 \text{Years Lived in Neighborhood}_i \\
 & + b_7 \text{Employed}_i + b_8 \text{Frequent Volunteer} + b_9 \text{PTA}_i + b_{10} \text{Married}_i \\
 & + b_{11} \text{Church Attendance} + b_{12} \text{Number of Organizations}_i \\
 & + b_{13} \text{Number of Dissats}_i + b_{14} \text{DCPS Grade} + b_{15} \text{Years of Education}_i \\
 & + b_{16} \text{Years of Education}^2_i + b_{17} \text{Years Child in School}_i + b_{18} \text{Grade Level}_i \\
 & + e_{ig}
 \end{aligned} \tag{9.1}$$

where $e_{ig} \sim N(0, S^2)$.

In this first cut, the assigned grades are treated as continuous (0-4) variables. To handle missing data due to item nonresponse, we employ a multiple imputation approach (King et al. 2001; Little and Rubin 1987; Rubin 1987; Little and Rubin 2002; Allison, 2002). Specifically, we impute five complete datasets of 995 observations each using a predictive mean matching model (Little 1988; Allison 2002: 59-63; Van Buuren et al. 1999) and average the results of the analyses using Rubin's method (Rubin 1987; Little and Rubin 2002: 210-12). One particular advantage of Rubin's partially parametric predictive mean matching approach over other techniques used in the applied literature is the restriction of imputed values to those observed in the sample. The results of the models using this imputed data set are presented in Table 9.1.

Table 9.1 Here

As Table 9.1 shows, the observed effect of charter school enrollment on parent satisfaction remains strong in this simple linear regression. We find the coefficients on charter enrollment (which range in magnitude from .28 for principals to .54 for facilities) to be positive and statistically significant (all $p < .01$) on all four parental evaluation measures.

Controlling for Self-Selection Using Propensity Score Matching

There are several limitations of this simple linear approach that we must consider before concluding that our data support the hypothesis that charter schools *cause* an increase in satisfaction. First, we are assuming that our model is correctly specified—that the covariates predict the outcome in a linear, additive function and that our assumption of normality of the dependent variables is appropriate. Second, we are assuming that there is no endogeneity among the outcomes and their predictors. Finally, we are

assuming that there are no omitted variables that are correlated with both charter enrollment and our satisfaction outcomes; it is well-known that results from quasi-experimental studies of the effects of public policy (or other “treatments”) are potentially biased when the factors predicting self-selection into the program (here, charter schools) are correlated with the outcome measures.

A variety of techniques have been developed to deal with this last cause of bias. One solution is the estimation of some form of “Heckman-type” parametric treatment effects model, usually by means of a consistent two-step or full-information maximum likelihood model that relies on assumptions about the error structure for identification (Maddala 1983). Instead, in our analysis we use a semiparametric estimator, propensity score matching, originally introduced by Rosenbaum and Rubin (Rosenbaum and Rubin 1983; Rosenbaum and Rubin 1985) in a biometric context and increasingly used in econometric studies, ranging from evaluating the effects of training programs on subsequent earnings (e.g., Bryson, Dorsett, and Purdon 2002; Dehejia and Wahba 1999; 2002; Heckman, Ichimura, and Todd 1997) to evaluating successful techniques of heart catheterization (Hirano and Imbens 2001) to evaluating education policy (Schneider and Buckley 2003).

The logic underlying this method is to construct, from quasi-experimental data, a new variable, the propensity score, that summarizes pre-treatment (and other potentially confounding) characteristics for each respondent. This score is essentially the conditional probability of the subject selecting the treatment. Based on these propensity scores, a new sample consisting of the treatment group and their matched controls is created and the size and significance of the treatment effect can be estimated using this sample.

Propensity score matching has several advantages over the Heckman-type treatment effects models, such as the relaxation of restrictive parametric assumptions and of the need to find instrumental variables for practical model identification (see LaLonde 1986 and Puhani 2000 for a discussion of the sensitivity of Heckman-type of models to misspecification and to violations of their parametric assumptions). Moreover, as Dehejia and Wahba (1999) argue, matching provides estimates of the treatment effects more similar to randomized field trials than can be obtained using other corrections for self-selection.

However, the propensity score model produces correct causal inference only if selection to treatment is independent of outcomes, conditional on the covariates used in the matching procedure. Thus there is a tradeoff: the Heckman-type modeling strategy relaxes the assumption of observability of all covariates correlated with both the treatment decision and the outcome, but imposes a strong assumption about the error structure. Conversely, the more intuitive propensity score matching approach assumes selection on the observables, but does not require strict parametric assumptions for estimation. Our strategy is use the propensity score approach combined with a test of sensitivity to hidden bias introduced by Rosenbaum (2002 [1995]: 105-138) and discussed by DiPrete and Gangl (2004) in the context of unemployment insurance.

Our approach is thus superior to the classical linear regression model both in adjusting on the observable covariates without strong assumptions and in testing for hidden bias. However, we have not considered yet the issue of endogeneity. Our strategy is to take advantage of the longitudinal nature of the “treatment” (that is, enrollment in charter schools) to better disentangle cause from effect.

To do that, in the next analysis we select the 226 (137 charter, 89 DCPS) parents whose child was new to their school in Fall 2001, the time of the first panel interview. We match these new charter parents to the new DCPS parents (sampling with replacement) on their estimated propensity scores, and then consider the differences between these matched pairs using the grades assigned by parents in the Spring of 2001, the second panel wave. This method of analysis allows us to draw more valid causal inference about the effects of charter school enrollment by correcting for self-selection factors *prior* to the treatment but considering outcomes one school-year later. The trade-off is a potential reduction in power to detect effects due to the decrease in sample size caused by our restriction of the sample to students new to their school.

We assume that all missing data, including panel attrition, is missing at random (MAR) conditioned on the observed covariates, so we impute these missing values using the same predictive mean matching method discussed above.

To implement this propensity score matching procedure, we first estimate a probit model (other generalized linear models or semiparametric models for dichotomous data would also be appropriate) to generate each individual's predicted probability of choosing a charter school (that is, the propensity score). The covariates we use to estimate the propensity score are a subset of the covariates discussed in the regression model above. If C_i is the indicator of treatment (charter enrollment) for parent i and \hat{p} is the estimated propensity score for a vector, \mathbf{w}_i , of predictors, the full model is thus:

$$\begin{aligned}
\hat{p}(C_i = 1 | \mathbf{w}_i) = & \Phi(b_0 + b_1 \text{White}_i + b_2 \text{Hispanic}_i + b_3 \text{Other Race}_i \\
& + b_4 \text{Years Lived in D.C.} + b_5 \text{Years Lived in Neighborhood}_i \\
& + b_6 \text{Employed}_i + b_7 \text{Frequent Volunteer}_i + b_8 \text{PTA} \\
& + b_9 \text{Married}_i + b_{10} \text{Church Attendance}_i \\
& + b_{11} \text{Number of Organizations}_i + b_{12} \text{Number of Discussion}_i \\
& + b_{13} \text{DCPS Grad}_i + b_{14} \text{Years of Education}_i \\
& + b_{15} \text{Years of Education}_i^2)
\end{aligned}
\tag{9.2}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function. The results of the propensity score estimation model are presented in Table 9.2.

Table 9.2 Here

Once the predicted probabilities from this model are estimated, we match each charter parent to a single DCPS parent who is their “nearest neighbor” on the propensity score metric.⁷ We then verify that the balancing property is satisfied by examining the first and second moments of the distribution of propensity scores for treated and non-treated units. This property ensures that the pretreatment characteristics of the respondents within equally spaced intervals of the propensity score, independent of their treatment status, have the same mean and variance (Rosenbaum 2002: 295-328; Becker and Ichino 2002).

To test the hypotheses that the charter parents have higher levels of satisfaction reflected in the grades they assign, we use the nonparametric Wilcoxon signed-rank test for the matched pairs, a nonparametric test which has the added advantage of allowing us to relax the assumption that the outcome measures are continuous and normal. Since the

⁷ We use the one-to-one matching algorithm in the software provided by Leuven and Siansi (2003). Note also that in this analysis we are assuming a simple, additive treatment effect and that the stable unit treatment value assumption (SUTVA—that the treatment status of one unit does not effect the assignment to treatment of other units).

Wilcoxon procedure tests the null of equality of samples but does not produce a point estimate, we also compute the Hodges-Lehmann point estimates of the difference between charter and DPCS parents on each measure (Rosenbaum 2002 [1995]: 47-50).

As before, we average the results of the point estimates over the five imputed datasets to account for missing values. We also produce averaged p -values for the Wilcoxon tests by averaging their approximate z scores (under the large sample assumption of the signed-rank test) following Li et al. (1991). Finally, we also compute a multivariate satisfaction variable using the technique suggested by Rosenbaum (2002 [1985]: 50-53) for partially-ordered outcomes that enables us to consider the satisfaction results on all four dimensions simultaneously.⁸ The results of these analyses are presented in Table 9.3.

Table 9.3 Here

The first column of Table 9.3 provides the averaged Hodges-Lehman point estimates and the second column gives the averaged p -values of the signed-rank test. All four of the outcome measures, as well as the multivariate combination of them, are significantly larger in the charter group than in the matched control group, and the point estimates of the difference are similar in size to the estimates from the regression models in Table 9.1.

The final column of Table 9.3 presents the results of our sensitivity test for hidden bias discussed above. Rosenbaum's Γ (2002 [1995]: 110-119) is the ratio of the odds that the two cases (treatment and control), matched on the propensity score computed from

⁸ We construct this measure by comparing each matched pair on each of the four outcome measures. If the charter parent has a higher response than the control, we add 1 to the measure; we subtract 1 if the control parent's outcome is larger. The final measure thus ranges between -4 and 4.

observable data, will select the treatment. Here, we present the estimated value for Γ at which the signed-rank test is no longer significant at the .05 level. Thus for the case of the overall school grade outcome, for example, if we hypothesize that there is an unmeasured covariate correlated with both selection to charter schools and the outcome, this covariate would have to be of sufficient size to increase the odds ratio of selecting treatment by 1.64. That is, the charter parent would have to be 1.64 times more likely to choose the charter school due to this hidden bias in order for the Wilcoxon test to fail to reject the null of equivalent samples at the .05 level. As the table shows, we find values of Γ ranging from 1.29 (principal grade) to 1.91 (facilities). These values may appear fairly small, suggesting that hidden bias, if it exists, may be able to explain at least some of the results. However, to place them in a social science research context, they are all larger than the reported critical Γ 's in DiPrete and Gangl's (2004) study of the behavioral effects of unemployment insurance. Also it should be emphasized that this test does not allow one to reach conclusions about the *existence* of such bias—only to assess the sensitivity of the results in the case where such an unmeasured covariate exists.

We believe that these statistical techniques confirm that charter school parents evaluate their schools more highly than DCPS parents and that the observed higher grades are not the result of self-selection. We now turn to one further method for confirming this conclusion.

Judging the Value Added of Charter Schools:

The Results of a Natural Experiment

Randomized field trials are often held up as the “gold standard” for policy evaluation. In a well-known series of studies of the effects of school voucher programs

on academic achievement, Peterson and his colleagues (see, e.g., Greene 2002; Howell and Peterson 2000; Peterson et al. 2002) randomly assigned parents who have applied for vouchers to a “treatment condition” in which parents are given vouchers and to a “control condition” in which parents who sought vouchers are denied them. Since all the parents in these experiments sought vouchers and their assignment to the treatment and experimental group was randomized, the factors that contaminate so much social science evaluation work, including differences in motivation to seek the “treatment,” are ideally orthogonal to the outcomes by design.⁹

Peterson’s ability to even attempt random assignment was the result of a unique set of conditions, including the fact that the voucher programs he studied were privately financed. Most education reforms do not present researchers this type of opportunity, and given that charter schools are publicly financed and popular among parents and policy makers alike, a carefully constructed randomized field trial has up-to-now been difficult to implement.¹⁰ However, since the demand for charter schools in Washington DC exceeds supply and that charter schools facing excess demand conduct lotteries, there is a population whose motivation to choose charter schools is at the same level of charter school parents (that is, both sets of parents have chosen to choose), but who through the luck of the draw were denied access to charter schools. In our random sample of parents whose child was in the traditional DC public schools approximately 13 percent of the parents (60 in all) said that they had tried to enroll their child in a charter school but

⁹ There have been well-known and very intense debates about how well this design has been implemented and about the effects of vouchers (e.g. Kruger and Zhu 2004; Barnard et al. 2003). In addition, others have criticized randomized trials as introducing new biases and complication to policy evaluation (e.g. Hanushek 1999 on class size experiments in education and Heckman and Smith 1995 on “randomization bias” in social experiments in general).

¹⁰ As noted earlier, as of this writing, the U.S. Department of Education is in the process of launching such a study. However, the results of this study will not be available for several years.

failed. While not as elegant as a fully controlled experiment, we believe that the comparison of charter school parents (who received the “treatment”) and these other parents who sought the treatment but were denied access presents yet another way of assessing the sensitivity of our results to hidden bias.

As in the propensity score models, are outcome measures are the grades assigned in Spring 2002 by charter parents whose children have completed their first year in the school. Here, however, we compare them to the Spring 2002 evaluation measures of all 60 “denied access parents” (DAPs), regardless if their child has completed their first year or if they have been in the school longer. Partly we choose this comparison group due to the small sample size, but we also believe it to be a theoretical appropriate group, since both groups shared the motivation to try to get into charter schools, but some of them were denied enrollment in the charter schools at random via a lottery.

We analyze the data using the Wilcoxon rank-sum test for the univariate outcomes and Leach’s (1991) multivariate nonparametric test for two samples for multivariate satisfaction. Once again, we average these statistics over all five multiple imputation datasets (Li et al. 1991). We present the results of these analyses in Table 9.4.

Table 9.4 Here

Here our results are somewhat different than in the regression or propensity score models presented above. The charter sample’s mean outcomes are higher for all four univariate measures, but we find a statistically significant difference between the charter and DAP parents only for the facilities and overall school grades. Note that the two measures for which the Wilcoxon rank-sum test fails to reject the null hypothesis—the evaluation of teachers and principals—are also the measures which had the smallest

critical values of Γ in the propensity score sensitivity analysis (i.e. the results potentially most sensitive to bias due to an omitted variable). The multivariate comparison suggests that there is still a significant difference between charter and DAP parents overall ($p = .025$). Since it is possible that the facilities effect is strongly affecting the overall school grade, we also conduct a multivariate comparison of the two samples omitting the school grade measure, with similar results.

While it is clear from our data that charter parents on average evaluate their child's school more highly than their counterparts in the traditional public schools in D.C., the results of the DAP comparison raise questions about the charter effect in the areas of teacher and principal satisfaction. It is possible that the small sample size of the DAPs, combined with additional uncertainty caused by the missing data imputation, has left us with insufficient power to detect the difference at conventional levels of significance. It is also possible that the DAPs are not perfect counterfactual controls to the charter parents—either because the lotteries did not perfectly randomize the groups or because of dissatisfaction or “randomization bias” (Heckman and Smith 1995) induced by their failure to win the school lottery.¹¹

We suspect that the truth is somewhere in between. Parents in charter schools are, in general, more satisfied after a single school year. However, this difference in satisfaction is sensitive to other factors, perhaps especially in the case of evaluating

¹¹ One of the least discussed aspects of many school choice programs is the extent to which *schools choose students* as compared to the more frequently asked question of how *students choose schools*. There is some anecdotal evidence that more than the luck of the draw is involved in the selection process. For example, some charter schools have been accused of announcing very short deadlines for applications (and then recruiting the students they want to apply before the deadline). When all these desired students are signed on, the school then extends the deadline (if there are seats left over) and conducts the mandated lottery. One consequence is that the parents who have lost the lottery may actually learn that the deck was stacked against them, causing them to evaluate the schools in which their children remain even more negatively. See also Fiske and Ladd (2000, 285-6). We discuss this point further in the concluding chapter.

teachers and principals. We return to this question in the next chapter, where we examine the persistence of this difference in satisfaction over time. Before doing so, however, we look at parental responses to four other important indicators of satisfaction.

Satisfaction with Other Dimensions of Schools

In the last section, we used grades to reflect parent evaluation with their child's school; however, most existing work has focused explicitly on reported levels of satisfaction with various characteristics of schools. Therefore, we queried parents specifically on their level of satisfaction with several other dimensions of the schools their children attended—each of which has important implications for the quality of schooling. Our empirical model here is identical to the second model for grades discussed above—we use estimated propensity scores of parents with children new to their schools to match charter parents to their DCPS controls and then measure the outcome variables at the conclusion of the school year.

Satisfaction with Values

One of the more interesting findings in this chapter has to do with parental satisfaction with the attention paid to values in their child's school.¹² A consistent theme in the research on effective schools is the importance of values in creating a strong community around which effective schools can be built. This theme is particularly evident in the analysis of the success of Catholic schools, where many scholars believe that the communal nature of these schools, built on a shared sense of values, helps

¹² We asked parents how satisfied they were to the attention of the school to “values” without defining the term. C.S. Lewis, in his *Studies in Words*, identified “the tendency of words to become less descriptive and more evaluative . . . and to end up by being purely evaluative—useless synonyms for good and bad.” We believe that the term “values” clearly central to many political debates (including that over schools) has that evaluative component. In turn, we know that the term is essentially undefined—but has strong emotional connotations, as evidenced in the 2004 Presidential election campaign.

produce academic success (Bryk et al. 1993; Coleman et al. 1982; Hoffer et al. 1985).

While there is not much research on the extent to which *public* schools of choice, including charter schools, create a commitment to values, we believe that charter schools also have the potential for using shared values to create stronger communities and through such stronger communities, greater academic success. While we explore the issue of the strength of school communities in greater detail later in this book (Chapters 11 and 12), in Table 9.5 we see that charter school parents are significantly more satisfied with the attention to values in their schools compared to DCPS parents.

Satisfaction with Discipline

We turn next to satisfaction with the discipline in the school. Unfortunately, urban public schools are all too often not safe and secure learning environments and inner city parents often seek out safe schools for their children. To measure satisfaction with discipline, often one of the selling points of charter schools to parents in D.C., we asked parents how satisfied they were with the discipline in their child's school using a five point scale. The results in Table 9.5 below suggest that parents in charter schools in our cross-section are *not* more satisfied with the level of discipline than are parents in the traditional DC public schools.

We also asked parents how satisfied were they with two other aspects of schools that have been associated with higher academic performance: school size and class size.

Table 9.5 Here

Satisfaction with Size of School

There is an increasing body of evidence that suggests that small schools are generally better than large ones and that the benefits of small schools are particularly

pronounced in enhancing student achievement in lower income communities (Fowler and Walberg 1991; Duke and Trautvetter; Cotton 1996; Lee and Smith 1997; Howley, Strange, and Bickely 1999).

While there are many reasons small schools may work better, from our perspective Wasley et al. (2000) make the most important link. They argue that small schools improve education by creating intimate learning communities in which students are well-known to each other and to their teachers and are encouraged by adults who care for them and about them. These smaller, more intense communities, in turn, reduce the isolation that adversely affects many students; reduce discrepancies in the achievement gap that plagues poor children; and encourage teachers to be more creative in their ways of thinking and teaching styles.

Small schools have also been linked to higher levels of cooperation between teachers, better relations between teachers and school administrators, and more positive attitudes towards teaching (see, for example, Stockard and Mayberry 1992; Lee and Loeb 2000). In addition, small schools often foster parental involvement, which benefits students and the entire community (see Nathan and Febey 2001, also Schneider, Teske, and Marshall 2000).¹³

In short there is a wide range of benefits that accrue to smaller schools—and many of these benefits flow from fundamentally shifting the nature of the school community. As the reader will recall from Chapter 3, at the time these data were collected, charter schools in DC were, on average, smaller than the traditional DC public schools. Charter elementary schools averaged 271 students, compared to almost 450

¹³ Cotton 2001 presents an excellent review of the literature examining the effects of small learning communities.

students for DC public schools and charter middle schools averaged 354 students compared to 543 students for DC public schools. The difference at the high school level was even greater: around 150 for charter high schools compared to around 600 students for DC public high schools.¹⁴ Given these large objective differences, it is probably not surprising that charter school parents are more satisfied with the size of their child's schools than are other DCPS parents.

Satisfaction with the Class Size

While the research linking school size and desired outcomes has accumulated in a relatively straightforward manner, the research linking class size to learning outcomes has been much more contentious—but there is a growing consensus that students in smaller classes do better.

Some researchers have argued that educational inputs, including class size, are not associated with higher performance. While economist Rick Hanushek has been a leader staking out this position (see, e.g., Hanushek 1997), other researchers using a range of data also have also found that reducing class size has no effect on educational outcomes (see, for example, Hoxby 2000 or Johnson 2000) or can have unintended consequences in the distribution of teachers across schools of varying socioeconomic status (Bohrstedt and Stecher 2000). But this work, and Hanushek's in particular, has been subject to criticism (see, for example, Greenwald, Hedges, and Laine 1996 and Krueger 2000) and there is a growing body of empirical work documenting the importance of smaller classes (for example, Ferguson 1991; Folger and Breda 1989; Ferguson and Ladd 1996; or Wenglinsky 1997).

¹⁴ These differences will most likely not endure over time. Many of the differences are the result of the newness of charter schools, many of which are just ramping up to full size, often by adding grade levels.

While the econometric evidence has been intensively fought over, there have been a series of experiments in which class sizes have been reduced, and the results of these experiments clearly support the benefits of smaller class size. Studies in a number of states have found positive results of smaller class size (see for example Egelson et al. 1996; Molnar et al. 1999), but the STAR program, authorized by the Tennessee legislature in 1985, has received the most attention.

The results from the STAR experiment have shown significant positive effects, especially for lower SES students and for African-American students (see for example, Mosteller 1995 and work of Krueger (e.g., Krueger 2000 and Krueger and Whitmore 2000)). While the debate about small class size continues among researchers, parents, educators, and policy makers, have already made their preferences clear: throughout the nation small class size is one of the most popular school reforms.¹⁵

While we have no objective data available about average class size in the DC schools, we do see from Table 9.5 that parents in charter schools are significantly more satisfied with their child's class size than are parents with children in the traditional public schools.

Are Charter School Parents Wearing Rose-Colored Glasses?

Parents in charter schools evaluate their child's school more highly and are more satisfied with many dimensions of those schools than parents with children in traditional DC public schools. These differences survive a wide range of tests to control for possible biases introduced by self-selection built into the option demand nature of charter schools.

¹⁵ The preference for smaller classes is being codified in law: nearly half the states have enacted legislation and are spending hundreds of millions of dollars each year to reconfigure school buildings to reduce the student-teacher ratio to twenty or fewer students per teacher (Magnuson 2000).

In the next section of this chapter, we turn to another source of potential bias that might be inflating the evaluations of charter school parents.

In our earlier discussion of behavioral decision theory, we briefly mentioned the theory of motivated reasoning: When people choose a program or a service, they have a psychological bias towards evaluating it positively. In the literature on school choice, this is sometimes referred to as “rose-colored glasses” or a “rose-colored glasses effect” (Erikson 1982; Erikson 1986; Goldring and Shapira 1993).

If true, this phenomenon complicates the task of the analyst using self-reported levels of satisfaction from attitude surveys to evaluate the perceived quality of schools. While virtually every study of schools of choice finds higher levels of satisfaction among participating parents and students versus their counterparts in the traditional public schools (e.g., Bridge and Blackman 1978; Peterson 1998; Schneider et al. 2000; Moe 2001; Schneider and Buckley 2003), the possibility of the rose-colored glasses effect calls these results into question.

To try to untangle this problem, we treat the rose-colored glasses effect as a measurement issue. We compare traditional public school and charter school parents on three measures of school satisfaction employing a new model, the CHOPIT (Compound Hierarchical Ordered ProBIT) model designed explicitly to aid applied empirical research comparing across groups using ordinal survey responses (King et al. 2003).

We will show that choice alone does *not* appear to systematically bias perception. In fact, we will show that the demographic characteristics of choosers tend to predict bias in the opposite direction: the type of parents who are dissatisfied enough to change their child’s school also tend to be tougher critics of the new school as well.

Rose Colored Glasses as a Case of Differential Item Functioning

The central assumption of our approach is that the rose-colored glasses effect is closely related to what is known in the educational testing literature as differential item functioning or DIF (Holland and Wainer 1993). From this perspective, the rose-colored glasses effect implies that parents in charter schools (or other schools of choice) who have the same true levels of satisfaction with their child’s school respond to evaluation questions with a different set of response probabilities than those in traditional public schools—the very act of choosing creates a “positivity” bias.

For ordinal responses, the act of choosing can be viewed as acting to shift the “cutpoints” that partition a hypothetical continuous latent variable into the measured discrete ordinal categories, rather than acting as an influence solely on the latent continuous random measure of satisfaction or evaluation. The CHOPIT model is designed to account for precisely this problem and to produce estimates not only of the outcome model of interest, but also the effect of a set of covariates that are hypothesized to affect these cutpoints (King et al. 2003).¹⁶

Central to CHOPIT is the use of “anchoring vignettes”—hypothetical cases designed by the researcher with a known relative ranking on a dependent variable of interest. In our study, this meant presenting parents with descriptions of hypothetical schools that are fixed *a priori* to have a known and transitive ranking (for example, a best, middle, and worst school on a particular dimension such as parent-principal communication). The respondents are asked to evaluate the anchoring vignettes and these

¹⁶ Appendix 9.1 gives more details on the CHOPIT model.

evaluations are then used to properly scale the true question of interest (referred to as the “self-assessment question”).

More precisely, in the context of the familiar ordered probit model with a latent continuous variable (Zavoina and McElvey 1975), CHOPIT allows the researcher to estimate the “cutpoints” that transform the underlying latent variable to the observed categorical response (and the contribution of a vector of covariates to those cutpoints) simultaneously with the mean model that is usually estimated. It is thus possible to model the potential differential item functioning and properly compare (with cutpoints estimated at the individual level, if data allow) categorical responses.

The rose-colored glasses effect, in our CHOPIT model, is estimated by the contribution of an indicator variable for having a child in a DC charter school as opposed to a traditional public school to the estimation of the cutpoints of the mean evaluation model, holding the effects of demographics and other attitudes constant.

Anchoring Vignettes

To implement the CHOPIT model, we asked parents to evaluate their child’s school on three different dimensions (parent-parent relations, parent-teacher relations, and parent-principal relations), and then to evaluate three hypothetical schools on each dimension.¹⁷ We present parent-principal relations as an example.

¹⁷ The subset of data used in this analysis is from the second wave of the telephone survey, with a sample size of 374 (after loss due to panel attrition and listwise deletion of missing responses). Two hundred of the parents in the sample have a child in a charter school.

We begin with the self-assessment question; in this example asking the parent to grade the relationship between the parent and the principal in their child's school on the same familiar A, B..., F scale.¹⁸ We then presented the following anchoring vignettes:

Anchoring Vignettes: Parent-Principal Relations

Let's consider this situation: there is a child who routinely disrupts class and gets sent to the principal's office on a regular basis.

1. At this school, the principal calls the parents, follows up with a letter, and actively tries to find out the source of the child's behavior problem. How would you grade the relationship between the principal and the parents at this school?
2. At this school, the principal sends a letter home notifying the parents of the problem but doesn't do anything else. How would you grade the relationship between the parents and the principal at this school?
3. At this school, the principal just has the child sit in the main office for a few hours and never notifies the parent. How would you grade the relationship between the principal and the parents at this school?¹⁹

If charter parents are viewing their child's school through rose-colored glasses, then the coefficient on the charter variable in the equations for the higher-valued (more positive evaluation responses) cutpoints will be negative, while those of the lower-valued cutpoints may be positive, negative or indistinguishable from zero. For example, Figure 9.2 illustrates one possible scenario that is consistent with the rose-colored glasses effect. In the figure, which shows a hypothetical distribution of the latent variable for charter parents, the cutpoints between the grades of A and B and B and C are shifted to the left as a result of the effect of a negative coefficient on charter enrollment, while the other

¹⁸ These are the data we analyzed earlier in the chapter.

¹⁹ The order in which the three sets of vignettes was presented was randomized as was the order of the three alternatives.

cutpoints remain unchanged. Thus the probability of a parent reporting an A or a B for a given latent score is increased, while the probability of reporting a C is decreased.

Figure 9.2 Here

Does Choice Create Rose Colored Glasses?

The first model we estimate for each of the dependent variables is a simple model in which parents' responses are modeled using the same vectors of covariates for each of the three school community dimensions and the cutpoint models contain only a constant and an indicator variable for charter school enrollment. For the mean model in each dimension,

$$\begin{aligned}
 E(Y_i^*) = & b_1 \text{Charter}_i + b_2 \text{White}_i + b_3 \text{Hispanic}_i + b_4 \text{Other Race}_i + \\
 & b_5 \text{Education}_i + b_6 \text{Education}^2_i + b_7 \text{Married}_i + b_8 \text{Employed}_i + \\
 & b_9 \text{Church Attendance}_i + b_{10} \text{DC Public Schools Attended}_i + \\
 & b_{11} \text{Time Lived in DC}_i + b_{12} \text{Time Lived in Neighborhood}_i + \\
 & b_{13} \text{Female}_i
 \end{aligned}
 \tag{9.3}$$

and for the cutpoint model:

$$t_i^1 = g_0^1 + g_1^1 \text{Charter}_i
 \tag{9.4}$$

for the first cutpoint and,

$$t_i^k = t_i^{k-1} + \exp(g_0^k + g_1^k \text{Charter}_i)
 \tag{9.5}$$

for the subsequent ones, where Charter is a dichotomous variable indicating that the parent's child is in a charter school and the remaining covariates are as defined earlier in the chapter. The constant is omitted in the mean equation to allow for the estimation of all cutpoints.

Table 9.6 Here

Table 9.6 compares the results of the CHOPIT mean model to results from a typical ordered probit. The results are comparable, although note that the coefficients on charter enrollment are generally larger in the CHOPIT models (but are not readily interpretable given the nonlinearity of the model). Table 9.7 presents the results of the cutpoint models—the information we need to evaluate the presence of rose-colored glasses.

The model provides only slight support for the rose colored glasses effect: the only statistically significant coefficient on the charter indicator is the negative coefficient on the third cutpoint of the teacher-community model. Note, however, that because of the parameterization of the model, a change in one cutpoint has the effect of also shifting subsequent cutpoints—in this case the remaining cutpoint between A and B. In other words, choosing a charter school appears only to make parents more likely to give the teacher-parent relations in their school a grade of A or B (at the expense of C, as in the hypothetical example in Figure 9.2), holding all the effects of the mean model constant.

Table 9.7 Here

Before drawing conclusions about the rose-colored glasses effect on the basis of this evidence, we turn to a model of the cutpoints that includes additional covariates whose omission may be biasing the results presented in Table 9.7.

Table 9.8 presents the results of the cutpoint models. The first cutpoint is now modeled by:

$$t_i^1 = g_0^1 + g_1^1 \text{Charter}_i + g_2^1 \text{White}_i + g_3^1 \text{Hispanic}_i + g_4^1 \text{Other Race}_i + g_5^1 \text{Education} + g_6^1 \text{Education}^2 + g_7^1 \text{DC Public Schools Attitude}_i + g_8^1 \text{Female}_i \quad (9.6)$$

and each subsequent cutpoint by:

$$t_i^k = t_i^{k-1} + \exp(g_0^k + g_1^k \text{Charter}_i + g_2^k \text{White}_i + g_3^k \text{Hispanic}_i + g_4^k \text{Other Race}_i + g_5^k \text{Education}_i + g_6^k \text{Education}^2_i + g_7^k \text{DC Public Schools Attitude}_i + g_8^k \text{Female}_i) \quad (9.7)$$

where the covariates are the same as described above.

Table 9.8 Here

Table 9.8 compares the results of this second set of CHOPIT mean models to results from their ordered probit counterparts. Again, the results of the two models are comparable, although note that here the coefficient on charter enrollment is generally smaller in size (and in some cases no longer statistically significant at conventional levels) in the CHOPIT models. Table 9.9 presents the results of the revised cutpoint models (4) and (5). Here the result is clear: although a variety of demographic covariates have a statistically significant effect on the cutpoints or thresholds in the CHOPIT model, the coefficients on charter enrollment are statistically indistinguishable from zero. This suggests that the thresholds with which respondents partition their underlying attitude into categories is not influenced, *ceteris paribus*, by choice alone.

Table 9.9 Here

Is there any difference, then, in the evaluations of charter and traditional parents? The answer is found not from the marginal effect of charter enrollment, but instead in considering the differences between charter and traditional public school parents on the variables used in the cutpoint model presented in Table 9.8.

Table 9.10 Here

As Table 9.10 shows, charter and DCPS respondents in the sample are alike in many ways, but differ in their attitude toward the DC Public School system and in their likelihood of reporting their race as white. In order to see what effect these differences have on the cutpoints in the CHOPIT model, we compute predicted values for each t_i , using one vector of sample mean values on the demographic and DCPS Grade covariates for charter respondents, and a second vector of sample mean responses for DCPS parents. The results for all three outcome measures, presented in Figure 9.3, are quite interesting.

Figure 9.3 here

The relative heights in the figure are like different-sized “hurdles” that must be overcome by the latent evaluation variable in order for the next highest grade to be reported. In all three measures, the predicted values for the charter t_i ’s are larger than those of the DCPS parents (although the differences are not significant for the parent-parent measure). Because higher thresholds means a more positive latent attitude is required to report a given categorical grade level, charter parents, according to these results, are *tougher* graders.

The findings of the CHOPIT model, using community measures of hypothetical schools to anchor parental evaluation of actual schools, thus have an interesting interpretation: charter parents do not appear to wear rose-colored glasses as a function of choice; in fact, quite the opposite appears true considering the demographic and prior attitudinal differences between the groups of parents. Choosers may be more demanding consumers, even after they have invested the time and energy to change their child’s

school. Rather than viewing their schools through rose-colored glasses, the type of people who choose may actually be wearing “grey colored” ones.²⁰

Clearly, parental satisfaction is a serious consideration in evaluating the success of charter schools, but students are ultimately the prime “consumers” of schooling. We look next at how students evaluate their schools to assess the extent to which greater parental satisfaction with charter schools is mirrored by higher student satisfaction.

How Do Students Evaluate Their School?

Paralleling the method we used to gauge parental satisfaction, we began by asking students in our study to grade their teachers, their school’s principal, the school facility and the school overall. We treated responses as ordered dependent variables and used ordered probit to explore differences between students in the two sectors. In the ordered probit equation, we controlled for the students racial identification, their grade level, and the length of time they were in the school. Because church attendance had a significant effect in some of our analyses of parental attitudes and behavior and because others have found religiosity to affect attitudes toward choice and the functioning of schools (see, for example, Moe 2001), we control for the student’s frequency of church attendance. We also controlled for the number of students in the school, since school size may have a significant effect on a large number of student behaviors and attitudes. In this section analyzing grading patterns, we also included the overall grade that students assigned to the DC public schools.

²⁰ An interesting aside: we asked parents to grade other services, including the parks and the quality of garbage pickup. There was a difference in the pattern of evaluations, with charter school parent evaluations significantly lower than DCPS parents. For example, while 19% of DCPS parents assigned the DC parks a grade of A, less than 14% of charter parents did so. Similarly, 25% of DCPS parents assigned a grade of A to the quality of garbage pickup, higher than the 21% of charter parents who did so. The grey-colored glasses charter parents wear when it comes to schools may also be reflected in their evaluation of other municipal services as well.

To deal with the missing values problem, as earlier in this chapter, we employed multiple imputation and the results reported in the next section are based on that procedure. As is evident in Table 9.11, we find no charter school enrollment effect on the assignment of any of the four grades—while charter parents clearly grade their child’s school higher than their DCPS counterparts, student evaluations are not any different.

As with our analysis of the parent satisfaction data, we next move beyond the analysis of grades and asked students what they thought about other characteristics of their school.

First, we asked them how much they agreed with the following statements that pertain to the level of orderliness in the school:

- Rules for behavior are strict in my school.
- There is a lot of cheating in my school.
- I don't feel safe in my school.

We also asked student how satisfied they were with the level of discipline in their school and with their school’s attention to values.²¹

As evident in Table 9.12, and paralleling the data we report for school grades, there are no significant differences in the responses of DCPS students and charter school students. So while charter parents, for example, are far more satisfied with the attention to values paid by the school, for students, this (nor any other characteristics we measure) sets them apart.

²¹ For the agree/disagree questions, we read students a four point scale: agree strongly, agree somewhat, disagree somewhat, and disagree strongly. They had the option to say that they didn’t know or that they neither agreed nor disagreed. For the questions pertaining to satisfaction with discipline and values, we used a four point scale and we report the percent that were very satisfied. They also had the option of saying they were neither satisfied nor dissatisfied.

We turn next to two measures that reflect Hirschman’s well-known exit, voice, and loyalty options. First, there is no statistical difference in student interest in the exit option: for example, while 74% of traditional public school students disagreed with the sentence “I wish I could go to a different school”, this was not significantly higher than the 68% of charter school students who felt the same way.

We do find a charter school effect (at $p=.08$, two tailed) for how proud students are of their school. More specifically, 38% of charter school students either agreed or agreed strongly with the statement that “Students are proud to go to my school”, significantly higher than the 26% of the DCPS students who agreed.²²

Conclusions

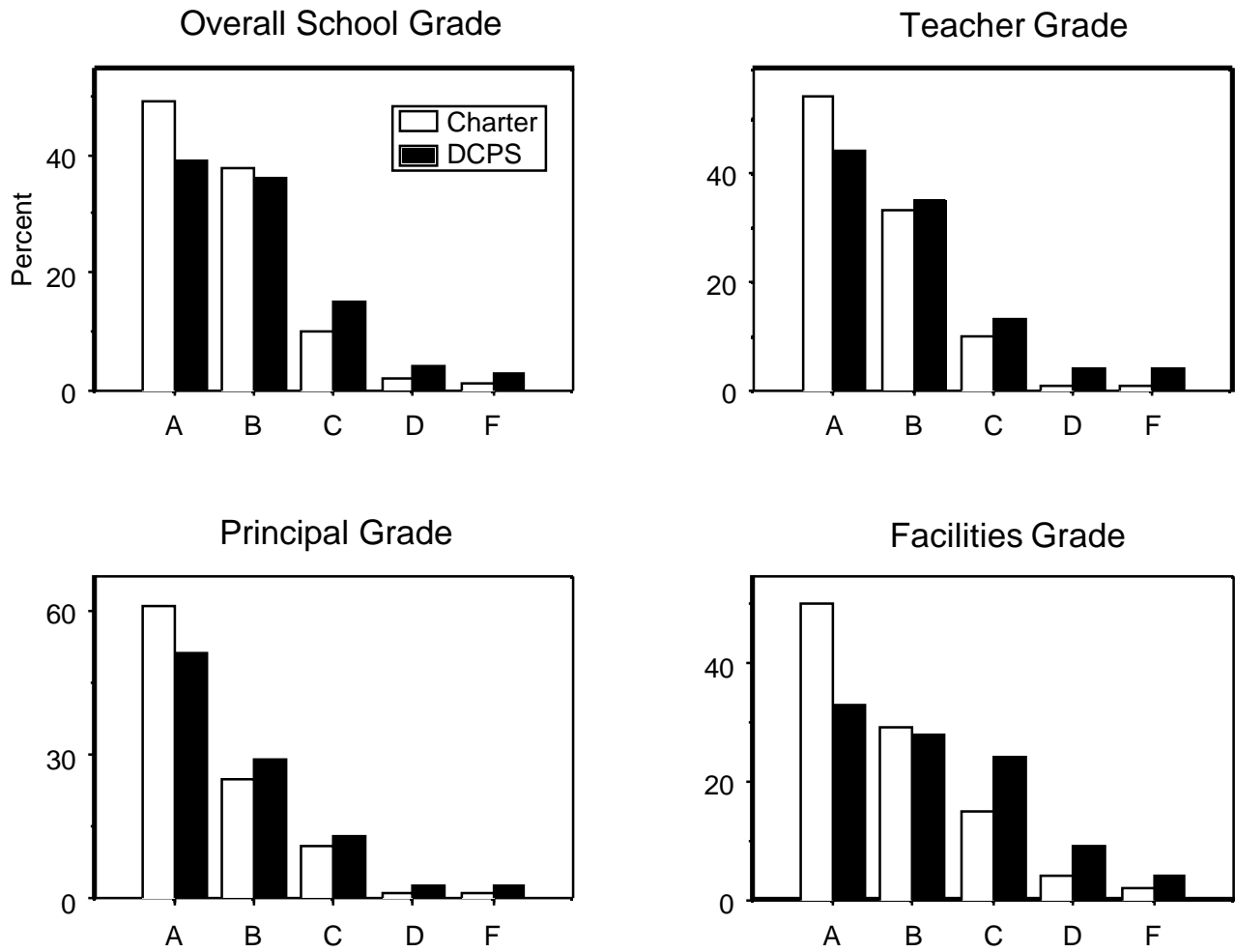
We began this chapter by confirming the findings that many other analysts studying schools of choice have previously found—parents who choose their children’s schools are more satisfied. However, we established a much stronger footing for these findings than is the norm, by explicitly addressing the possibility of bias emerging from self-selection and from rose-colored glasses. This is clearly good news for proponents of charter schools and, more generally, school choice—one of the most fundamental tests of the benefits of markets, greater consumer satisfaction, has been met. In addition, given the higher levels of satisfaction, charter parents may represent a political force for the further expansion of choice.

However, when we turned to the students who are the ultimate consumers of the product schools offer, we find far weaker results—students in both sectors are

²² As elsewhere in this book, these predicted response probability are generated via stochastic simulation holding all other covariates at their mean or modal values.

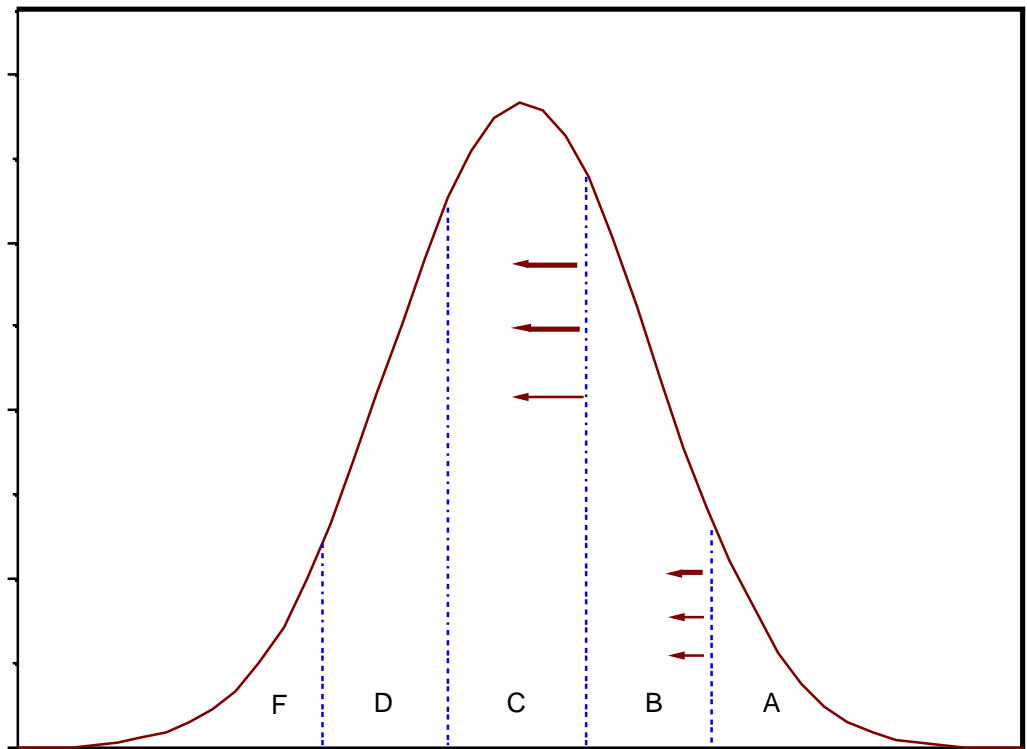
remarkably similar in their evaluations of their schools. The disjuncture between parent and student evaluations is an interesting phenomenon. In several other chapters we present later in this book, we will again see that charter school parents are more enthusiastic about the charter schools than are their children who actually attend them. We explore some potential reasons for this in our concluding chapter. But next, we turn to an analysis of the durability of these patterns over time.

Figure 9.1: In Wave 1, Charter School Parents Grade Their Schools Higher on Every Dimension



Note: Percentages are unadjusted responses from the wave 1 (Fall 2001) data, missing values deleted. Number of observations = 995.

Figure 9.2: How Rose Colored Glasses May Affect Evaluations



A possible hypothetical shift of cutpoints due to the rose-colored glasses effect. The combined results of the shifts is to make it more likely for a respondent with a given latent score to report an A or B grade, and less likely to report a C. Here the probabilities of reporting a D or F remain unchanged.

Figure 9.3: Charter Parents Have *Higher* Mean Thresholds Due to Demographic Differences

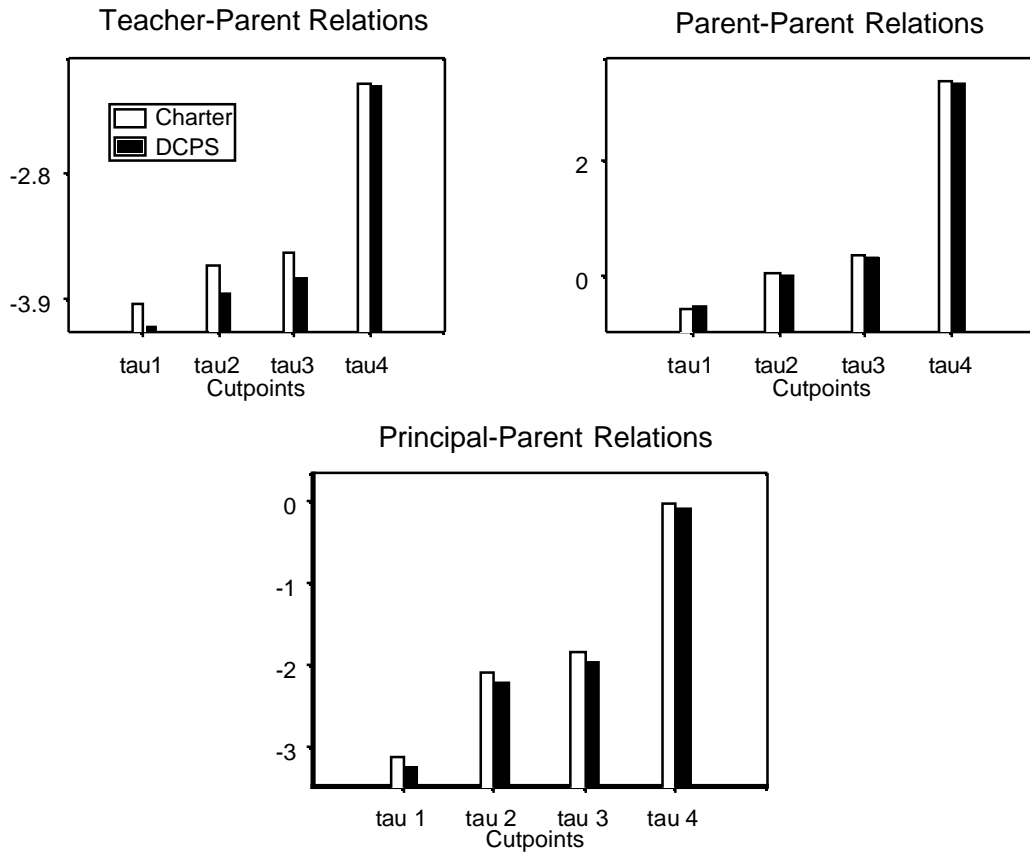


Table 9.1: A Simple Linear Model of Charter Effect

	School	Teacher	Principal	Facility
Charter	.41 (.07)	.35 (.07)	.28 (.07)	.54 (.08)
White	.06 (.14)	.14 (.14)	.27 (.16)	-.46 (.18)
Hispanic	.20 (.17)	-.05 (.17)	.18 (.13)	.13 (.19)
Other	-.18 (.17)	-.09 (.16)	.13 (.15)	.03 (.18)
Years in D.C.	.01 (.01)	-.02 (.01)	.01 (.01)	.0004 (.01)
Years in Neighborhood	.01 (.01)	.01 (.01)	.002 (.006)	-.005 (.007)
Employed	.03 (.10)	.05 (.10)	-.06 (.11)	-.11 (.11)
Frequent Volunteer	.08 (.09)	.21 (.08)	.27 (.10)	.05 (.10)
PTA Member	.13 (.09)	-.04 (.08)	-.12 (.09)	.01 (.10)
Married	.07 (.08)	-.06 (.08)	.03 (.09)	-.10 (.10)
Church Attendance	-.01 (.02)	-.001 (.02)	.03 (.02)	.01 (.02)
Number of Organizations	.01 (.04)	.01 (.03)	-.001 (.04)	.03 (.04)
Number of Discussants	-.01 (.01)	-.02 (.01)	-.01 (.02)	-.01 (.01)
DCPS Grade	.30 (.04)	.27 (.04)	.21 (.04)	.27 (.05)
Years of Education	-.10 (.05)	-.03 (.05)	-.03 (.05)	-.08 (.07)
Years of Education ²	.005 (.002)	.001 (.002)	.001 (.002)	.004 (.003)
Time in School	-.04 (.02)	-.03 (.03)	-.04 (.03)	.02 (.03)
Student's Grade Level	-.03 (.01)	-.03 (.01)	.01 (.01)	-.05 (.01)
Constant	2.87 (.37)	2.92 (.41)	2.54 (.35)	2.74 (.47)
Omnibus F (p)	6.30 (<.001)	6.81 (<.001)	3.39 (<.001)	6.74 (<.001)

Number of observations = 995 (555 charter, 440 DCPS). Results are coefficient estimates from independent linear regressions of outcome measures, weighted to correct for oversample of charter parents and averaged over five multiple imputation datasets constructed via predictive mean matching. Standard errors in parentheses, *p*-values less than .10, two-tailed, are in bold. The omnibus F-test reported in the last row is the test of the null hypothesis that all coefficients are jointly equal to zero.

Table 9.2: Predicting Who Chooses Charter Schools

	Coefficient (Standard Error)
White	-1.22 (.47)
Hispanic	-.68 (.34)
Other	.04 (.32)
Years in D.C.	-.02 (.02)
Years in Neighborhood	.02 (.01)
Employed	.06 (.20)
Frequent School Volunteer	-.18 (.20)
PTA Member	.27 (.21)
Married	-.10 (.18)
Church Attendance	.11 (.04)
Number of Organizations	-.01 (.11)
Number of Discussants	-.01 (.02)
DCPS Grade	-.24 (.09)
Years of Education	.13 (.13)
Years of Education Squared	-.01 (.01)
Constant	-1.27 (.88)

Number of observations = 226 (137 charter, 89 DCPS, sample limited to parents which children in new school). Results are estimates from maximum likelihood probit model of charter school enrollment, weighted to correct for oversample of charter parents and averaged over five multiple imputation datasets constructed via predictive mean matching. *p*-values less than .10, two-tailed, are in bold.

Table 9.3: Charter Parents Grade Their Schools Higher

	Hodges- Lehmann Point Estimate	Wilcoxon Signed- Rank Test p	Critical Level of Γ
School Grade	.5	<.001	1.64
Teacher Grade	.4	.004	1.41
Principal Grade	.7	<.001	1.29
Facilities Grade	.4	.007	1.91
Multivariate Satisfaction	1.1	<.001	1.72

Hodges-Lehmann point estimates and Wilcoxon p -values (two-tailed) are averaged over five datasets imputed via predictive mean matching. The Rosenbaum's Γ column provides the values of Γ for which the upper limit of the confidence interval on the p of the signed-rank test is exactly .05. Sample size varies between 135 and 136 matched pairs depending for different imputed datasets due to discarding of observations not on the common support of the estimated propensity score.

Table 9.4: Comparing Charter Parents to Parents Denied Access to Charter Schools

	Wilcoxon Rank- Sum Test p	Leach Two- Sample Test p
School Grade	.007	
Teacher Grade	.317	
Principal Grade	.624	
Facilities Grade	.001	
Multivariate Satisfaction		.025
Multivariate Satisfaction (School Grade Removed)		.028

Wilcoxon and Leach p -values are averaged using the method of Li et al. (1991) over five multiply-imputed datasets. Sample size is 197 (137 charter parents and 60 DAPs).

Table 9.5: Charter Parents Are More Satisfied on Some Other Dimensions

	Hodges- Lehmann Point Estimate	Wilcoxon Signed-Rank Test p	Critical Level of Γ
Discipline	0.0	.194	1.00
Values	.5	<.001	1.70
School Size	.6	<.001	2.10
Class Size	.6	<.001	1.69

Hodges-Lehmann point estimates and Wilcoxon p -values (two-tailed) are averaged over five datasets imputed via predictive mean matching. The Rosenbaum's Γ column provides the values of Γ for which the upper limit of the confidence interval on the p of the signed-rank test is exactly .05. Sample size varies between 135 and 136 matched pairs depending for different imputed datasets due to discarding of observations not on the common support of the estimated propensity score.

Table 9.6: Comparing Standard Ordered Probit to CHOPIT Estimates for Three School Community Measures, Mean Model (1), Cutpoint Models (2) and (3)

Variable	Parent-Teacher		Parent-Parent		Parent-Principal	
	Ordered Probit Coefficient	CHOPIT Coefficient	Ordered Probit Coefficient	CHOPIT Coefficient	Ordered Probit Coefficient	CHOPIT Coefficient
Charter	.19 (.11)	.23 (.15)	.23 (.11)	.28 (.14)	.19 (.11)	.18 (.14)
White	-.04 (.24)	-.10 (.26)	.37 (.24)	.37 (.25)	.09 (.24)	.16 (.25)
Hispanic	-.34 (.22)	-.46 (.24)	.10 (.22)	.18 (.24)	-.11 (.21)	-.18 (.24)
Other	.07 (.23)	-.13 (.27)	-.15 (.23)	-.08 (.25)	-.08 (.22)	-.24 (.26)
Education	-.35 (.14)	-.38 (.17)	-.29 (.11)	-.24 (.12)	-.31 (.13)	-.21 (.14)
Education ²	.01 (.01)	.01 (.006)	.01 (.004)	.01 (.004)	.01 (.004)	.01 (.01)
Married	.06 (.11)	.05 (.13)	-.03 (.11)	-.04 (.12)	.01 (.11)	-.02 (.12)
Working	-.24 (.13)	-.25 (.14)	-.31 (.12)	-.26 (.13)	-.19 (.12)	-.21 (.13)
Church	-.02 (.03)	-.01 (.03)	.02 (.03)	.02 (.03)	.01 (.03)	.02 (.03)
DCPS Grade	.02 (.05)	-.01 (.06)	.18 (.05)	.15 (.05)	.11 (.05)	.10 (.05)
Time in DC	-.01 (.01)	-.01 (.01)	.01 (.01)	.02 (.01)	.01 (.01)	.01 (.01)
Time in Neighborhood	<.01 (.01)	.01 (.01)	<.01 (.01)	<.01 (.01)	<.01 (.01)	-.01 (.01)
Female	-.02 (.14)	-.03 (.16)	.05 (.14)	.06 (.15)	-.05 (.14)	-.10 (.15)
Cutpoint 1	-4.60 (.98)		-2.90 (.79)		-3.46 (.92)	
Cutpoint 2	-4.07 (.98)		-2.48 (.78)		-3.04 (.92)	
Cutpoint 3	-3.39 (.97)		-1.58 (.78)		-2.41 (.92)	
Cutpoint 4	-2.46 (.97)		-.53 (.78)		-1.50 (.92)	
Vignette 1		-1.95 (1.19)		.17 (.85)		-.33 (.97)
Vignette 2		-3.24 (1.19)		-.92 (.85)		-2.45 (.97)
Vignette 3		-4.52 (1.20)		-2.55 (.85)		-3.28 (.98)
ln S		.02 (.07)		.09 (.06)		

Number of Observations = 374, Standard Errors in Parentheses.

Table 9.7: A Simple Model of the Cutpoints Provides Only Modest Support for the Rose-Colored Glasses Effect

	Cutpoint 1				Cutpoint 2		
	Teacher	Parent	Principal		Teacher	Parent	Principal
Charter	.147 (.115)	.052 (.107)	.119 (.084)		.194 (.154)	-.060 (.144)	-.089 (.119)
Constant	-5.17 (1.21)	-2.59 (.854)	-3.08 (.978)		-.522 (.137)	-.433 (.112)	-.403 (.104)
	Cutpoint 3				Cutpoint 4		
	Teacher	Parent	Principal		Teacher	Parent	Principal
Charter	-.211 (.109)	.009 (.097)	-.133 (.112)		-.105 (.093)	.049 (.087)	-.030 (.105)
Constant	-.119 (.095)	-.098 (.081)	-.274 (.093)		-.100 (.080)	-.060 (.075)	-.206 (.086)

N= 374. Standard errors in parentheses.

Table 9.8: Comparing Standard Ordered Probit to CHOPIT Estimates for Three School Community Measures, Cutpoint Models (4) and (5) with Additional Covariates

Variable	Parent-Teacher		Parent-Parent		Parent-Principal	
	Ordered Probit Coefficient	CHOPIT Coefficient	Ordered Probit Coefficient	CHOPIT Coefficient	Ordered Probit Coefficient	CHOPIT Coefficient
Charter	.19 (.11)	.15 (.13)	.23 (.11)	.23 (.13)	.19 (.11)	.07 (.13)
White	-.04 (.24)	-.29 (.29)	.37 (.24)	.08 (.28)	.09 (.24)	-.07 (.28)
Hispanic	-.34 (.22)	-.47 (.26)	.10 (.22)	-.02 (.26)	-.11 (.21)	-.05 (.25)
Other	.07 (.23)	.01 (.27)	-.15 (.23)	-.37 (.26)	-.08 (.22)	-.21 (.25)
Education	-.35 (.14)	-.13 (.15)	-.29 (.11)	-.17 (.14)	-.31 (.13)	-.04 (.14)
Education ²	.01 (.01)	.01 (.01)	.01 (.004)	.01 (.01)	.01 (.004)	<.01 (.01)
Married	.06 (.11)	.07 (.12)	-.03 (.11)	-.01 (.11)	.01 (.11)	.02 (.11)
Working	-.24 (.13)	-.24 (.13)	-.31 (.12)	-.26 (.12)	-.19 (.12)	-.18 (.12)
Church	-.02 (.03)	-.02 (.03)	.02 (.03)	.02 (.03)	.01 (.03)	.01 (.03)
DCPS Grade	.02 (.05)	-.01 (.06)	.18 (.05)	.11 (.06)	.11 (.05)	.04 (.06)
Time in DC	-.01 (.01)	-.01 (.01)	.01 (.01)	.01 (.01)	.01 (.01)	.01 (.01)
Time in Neighborhood	<.01 (.01)	-.01 (.01)	<.01 (.01)	<.01 (.01)	<.01 (.01)	<.01 (.01)
Female	-.02 (.14)	<-.01 (.17)	.05 (.14)	.08 (.16)	-.05 (.14)	.12 (.16)
Cutpoint 1	-4.60 (.98)		-2.90 (.79)		-3.46 (.92)	
Cutpoint 2	-4.07 (.98)		-2.48 (.78)		-3.04 (.92)	
Cutpoint 3	-3.39 (.97)		-1.58 (.78)		-2.41 (.92)	
Cutpoint 4	-2.46 (.97)		-.53 (.78)		-1.50 (.92)	
Vignette 1		-.17 (1.10)		.44 (1.01)		.85 (1.03)
Vignette 2		-1.47(1.10)		-.60 (1.01)		-1.15 (1.02)
Vignette 3		-2.79 (1.10)		-2.23(1.02)		-1.99 (1.02)
ln S		.03 (.06)		.05 (.05)		-.09 (.06)

Number of Observations = 374, Standard Errors in Parentheses

Table 9.9: Choice Alone Does Not Lead to Biased Evaluations

	Cutpoint 1			Cutpoint 2			
	Teacher	Parent	Principal	Teacher	Parent	Principal	
Charter	.01 (.11)	.003 (.10)	-.06 (.08)	Charter	.07 (.14)	-.06 (.13)	-.02 (.12)
White	-.61 (.30)	-.19 (.21)	-.51 (.17)	White	.26 (.34)	-.18 (.30)	.45 (.21)
Hispanic	.33 (.20)	.30 (.19)	.20 (.15)	Hispanic	-.08 (.26)	-.37 (.30)	.06 (.22)
Other	.26 (.21)	-.23 (.21)	-.02 (.15)	Other	-.39 (.34)	-.11 (.26)	-.23 (.25)
DCPS Grade	-.06 (.05)	-.10 (.04)	-.03 (.03)	DCPS Grade	.01 (.06)	-.04 (.05)	-.02 (.05)
Female	.02 (.13)	-.03 (.12)	.18 (.09)	Female	.04 (.18)	.09 (.17)	.11 (.14)
Education	.15 (.13)	-.06 (.08)	.10 (.08)	Education	.79 (.30)	.90 (.28)	.69 (.22)
Education ²	-.01 (.01)	.01 (.01)	-.01 (.01)	Education ²	-.03 (.01)	-.03 (.01)	-.02 (.01)
Constant	-4.02 (1.32)	-1.50 (1.04)	-2.60 (1.05)	Constant	-6.22 (2.11)	-6.98 (2.02)	-5.62 (1.61)
	Cutpoint 3			Cutpoint 4			
	Teacher	Parent	Principal	Teacher	Parent	Principal	
Charter	-.09 (.10)	-.003 (.09)	-.07 (.11)	Charter	-.06 (.09)	.08 (.08)	-.02 (.10)
White	.13 (.20)	.09 (.20)	.29 (.21)	White	.10 (.17)	-.01 (.16)	-.39 (.24)
Hispanic	-.40 (.24)	-.50 (.25)	.03 (.22)	Hispanic	-.44 (.22)	.07 (.16)	-.57 (.25)
Other	-.22 (.24)	.03 (.18)	.12 (.20)	Other	.01 (.18)	-.18 (.18)	-.10 (.20)
DCPS Grade	.03 (.04)	.02 (.04)	-.04 (.05)	DCPS Grade	.01 (.04)	-.01 (.03)	-.02 (.04)
Female	-.11 (.12)	.04 (.12)	-.11 (.13)	Female	.02 (.11)	-.12 (.10)	-.06 (.12)
Education	.04 (.22)	.14 (.15)	.11 (.15)	Education	-.17 (.09)	-.11 (.08)	-.16 (.09)
Education ²	-.01 (.01)	-.01 (.01)	-.01 (.01)	Education ²	.01 (.003)	.01 (.003)	.01 (.003)
Constant	-.52 (1.52)	-1.14 (1.07)	-.79 (1.06)	Constant	.83 (.62)	.71 (.62)	1.01 (.65)

N= 374. Values in boldface are significant at least the .10 level, two-tailed. Standard errors in parentheses.

Table 9.10: Charter and DCPS Demographics in Sample

	<i>Mean for Charter Parents</i>	<i>Mean for DCPS Parents</i>
White	.010	.155
Hispanic	.060	.074
Other	.045	.063
DCPS Grade	1.77	2.18
Female	.84	.74
Education	13.77	14.00
Education Squared	194.66	205.09

Sample size = 374 complete cases, 200 charter and 174 DCPS.

Appendix 9.1: The CHOPIT Model

Mathematically, the CHOPIT model is straightforward (although in practice estimation can be time-consuming). To begin with, let Y_i^* denote the latent self-assessment measure (the parent's true grade for their child's school's parent-teacher relationship, for example). As in the case of a standard ordinal probit model, $Y_i^* \sim N(m_i, 1)$, where $m_i = X_i b$. The actual reported self-assessment, y_i is a choice of K ordered categories. The latent variable is assumed to be converted to the observed measure via a vector of thresholds, t_i such that $y_i = k$ if $t_i^{k-1} \leq Y_i \leq t_i^k, k = 1, \dots, K$, where $t_i^0 = -\infty$ and $t_i^K = \infty$. Note that the cutpoints are indexed over the individual observations—the crucial difference from the standard model. These individual-level cutpoints are themselves defined as a function of another vector of covariates V_i and parameters g . The first cutpoint for w_e is defined: $t_i^1 = g^1 V_i$, and the subsequent cutpoints are constrained to increase monotonically by assuming the functional form:

$$t_i^k = t_i^{k-1} + \exp(g^k V_i).$$

As discussed above the vignettes are used to identify (anchor) the cutpoints. Let q_j denote the mean response or evaluation by subjects (parents in the data discussed here) to vignette j . we assume normal variability in i 's perception of q_j , Z_{ij}^* , where

$$Z_{ij}^* \sim N(q_j, S^2). Z_{ij}^* \text{ is measured using a survey instrument with the same } K \text{ response}$$

options again partitioned by cutpoints so that the categorical response $z_{ij} = k$ if

$$t_{i1}^{k-1} \leq Z_{ij}^* \leq t_{i1}^k, k = 1, \dots, K. \text{ These cutpoints are determined using the same coefficient}$$

vector g as in the mean model and the same vector V_{we} of covariates. It is this assumption of identical g 's that identifies the model and allows the vignettes to be used to correct DIF at the individual level. The CHOPIT model, once specified, is then estimated via standard maximum-likelihood methods, where the likelihood function maximized is $L(b, \sigma^2, q, g | y, z, V, X)$.

In the King et al. (2003) paper, a more general model is presented that includes an individual level random effect in the mean equation. This is strongly identified only in the context of a latent variable analysis of several simultaneous self-assessment questions. Although it is possible to estimate this variant of the model in which the three dependent variables discussed here are treated as multiple measures of a latent school community strength factor, we choose not to here because we believe that the disaggregated models are of greater substantive interest to education policy analysts. We should also note that we estimate the CHOPIT models here using the *gllamm* user-written add-on for Stata 7.0 or above (Rabe-Hesketh and Skrondal 2002). The model may also be estimated using code for the *R* statistical environment that can be downloaded from <http://anchors.stanford.edu/>.