An Empirical Analysis of the Human Capital Dynamic Linkage from High School to Post-Secondary Education

Xuejuan Su
Economics, Finance and Legal Studies
the University of Alabama

July 2004

Abstract

How does students' high school performance affect their post-secondary education outcomes? This paper uses a multi-stage education model to identify this human capital dynamic linkage. Both instrumental variable approach and the maximum likelihood approach are adopted to control for the possible endogeneity of high school academic performance. It is found that knowledge and skills acquired at high school improves the learning efficiency and outcomes at post-secondary schooling. Hence besides the direct effect, an education program effective at high school also has the indirect effect lasting through post-secondary education. Conventional program evaluation ignoring the indirect effects may generate misleading policy implications.

Keywords: Human capital, Multi-stage education, Program evaluation
JEL classification: I20; C31

1. Introduction

This paper attempts to identify the human capital dynamic linkage from high school to post-secondary schooling. More specifically, we are interested in the cause-effect relationship of how students' high school academic performance affects their post-secondary education outcomes. Given the hierarchical structure of the education system, individuals have to finish lower stages successfully before attending higher stages, and the knowledge and skills acquired at lower stages improve the learning efficiency and outcomes at higher stages. Starting from Ben-Porath (1967) and followed by Lucas (1988), this has become a well accepted theoretical specification; yet the empirical literature has shed very little light beyond the positive correlation of academic performance across education stages. This paper uses a multi-stage education model to identify the causal relationship of the human capital dynamic linkage in the data.

Identifying the human capital dynamic linkage is essential in education program evaluation. Most, if not all programs cover only certain stages of the entire schooling process. Program intervention improves human capital output at a given stage by changing the relevant education inputs; this effect is direct and has been extensively analyzed in the existing literature. On the other hand, the improved human capital from the treated stage also acts as an input and improves the learning efficiency and outcomes at subsequent stages; this effect is indirect and has received only minimal attention in the existing literature. Ignoring the indirect effect in program evaluation may generate misleading policy implications.

This paper directly models the multi-stage education technology and identifies the human capital dynamic linkage from high school to post-secondary education. Instead of reducing the human capital output at a given stage down to all exogenous explanatory factors (including historical education inputs at previous stages), this multi-stage education technology treats human capital output from previous stages as an intermediate input in subsequent stages of education. Only with this multi-stage education model, are we able to address some of the previously overlooked issues. To what extent does existing human capital affects subsequent accumulation of human capital? How big is the indirect effect channeled through this human capital dynamic linkage? What are the corresponding policy implications, for example, early intervention programs versus later remedial programs? To my best knowledge, this paper is the first attempt to identify the human capital dynamic linkage and the indirect effect associated with it.

In this paper, the dataset used for estimation is NELS88, which covers two main stages of education: high school and post-secondary education. Since high school outcomes may be affected by unobserved factors that also affect post-secondary schooling outcomes, direct estimation without controlling for endogeneity leads to downward bias. When a linear model is used at the post-secondary schooling stages, the instrumental variable approach is adopted to control for endogeneity. Arguably high school quality has only direct impact on high school but not post-secondary schooling outcomes, so variables measuring high school quality are used as the instruments to control for endogeneity. Over-identification tests do not reject the choice of instruments, and Hausman tests suggest significant difference between ordinary least squares (OLS) and instrumental variable (IV) estimates. The human capital dynamic linkage is found significantly positive from high school to post-secondary schooling. Hence besides the direct effect, an education program effective at high school will also have the indirect effect at post-secondary education. When an ordered probit model is used at the post-secondary stage, the maximal likelihood approach (MLE) is adopted to jointly estimate the two stages with potential error correlation. The results share the same qualitative features of the linear model. Based on this human capital dynamic linkage, including the indirect effect into program evaluation may generate different policy recommendations.

The remainder of the paper is structured as follows. Section 2 compares and connects the current paper to the existing literature. Section 3 introduces the multi-stage education model and the data used for estimation. Estimation results when a linear model is used at the post-secondary stage (both OLS and IV) are reported in Section 4. As a robustness check, estimation results for an ordered probit model at the post-secondary stage (MLE) are reported in Section 5. Section 6 discusses the importance of the human capital dynamic linkage in program evaluation and policy recommendations and draws the conclusion.

2. Literature review

This paper has connections with several strands of the existing literature. As mentioned before, there exists a large literature on education program evaluation of the direct effect, i.e., how changes in education inputs at a given stage affect the education outcomes at that stage. Among many others, Card and Krueger (1992), Krueger (1999), Angrist and Lavy (1999, 2001) and Hoxby (2000) study the effect of schooling quality from various aspects like the class size and teacher quality. However, since the datasets commonly used are either cross-sectional or short panel that covers only the period when the program under study is effective, the indirect effect arising from the human capital dynamic linkage is totally ignored. Hence the analysis is essentially static.

A few exceptions exist, for example Currie and Thomas (1995, 2000, 2002), that attempt to analyze the long-term program effects with long survey datasets. However, without the multi-stage education specification to capture the dynamic relationship in human capital accumulation, those studies cannot disentangle the interaction between the initial program benefit and other education inputs in the long term dynamic setup. Currie and Thomas (2000) can only claim tentatively that "the effects of Head Start may fade out more rapidly among black students, at least in part because black Head Start children are more likely to subsequently attend inferior schools". This paper proposes the multi-stage education model and treats human capital output from an earlier stage as the intermediate input in the education technology at a later stage. Hence the parameters distinctly capture the human capital dynamic linkage and its interaction with other education inputs in the long term. It provides a general setup to solve a long standing question.

Another strand of the literature, the school transition models, is usually setup in a dynamic framework. Sequential choice models are used to capture the hierarchical nature of the educational system, i.e., only conditioning on finishing lower stages successfully, do individuals have the choices to attend higher stages. An early example is Mare (1980), and more recent work includes Cameron and Heckman (1998) and Todd and Wolpin (2003). Todd and Wolpin (2003) rightly emphasizes the interdependence and cumulative nature of sequential schooling decisions, and argues that the entire history of educational inputs is relevant in decision making. However, that study implicitly takes the reduced form interpretation. At any stage, the human capital dynamic linkage is omitted from the education technology, with its impact indirectly captured through the inclusion of previous education inputs.

This reduced form specification of the education technology has the advantage of estimation consistency, given that education inputs are exogenous. Yet by adopting the reduced form specification, we lose the power to identify the specific channel through which past education inputs affect current education outcomes. Different interpretations may be consistent with the same reduced form estimates. For example, recurring evidence shows that parental education level has significant explanatory power on children's academic performance in post-secondary schooling. This effect may arise mainly from parent-child interactions at early childhood, i.e., children with better-educated parents are better prepared for schooling, and the initial benefit is transmitted into later schooling. In this case, the indirect effect dominates. Or this effect may arise mainly from parent-child interactions at post-secondary schooling, i.e., college students with better-educated parents get more help on their course work and hence do better. In this case, the direct effect dominates. Policy implications are drastically different under the two situations.

Suppose the policy is to educate the parents when their children are in college, This policy would have very little impact on the children's outcomes in the first case, yet positive impact on the children's outcomes in the second case. The reduced form estimates cannot distinguish these two. These can be distinguished only with the multi-stage education model where human capital output from an earlier stage is an intermediate input in the education technology at a later stage.

Methodologically, this paper shares substantially with the literature on treatment effect. Treatment effect models usually take the form of a two-equation system, with one governing the treatment decision and the other governing the outcome. Unobserved factors that affect both the treatment decision and the outcome leads to the self selection bias, if endogeneity is not properly controlled. [1] Similar endogeneity problem may exist in the education process, where high school outcomes may be affected by unobserved factors that also affect post-secondary schooling outcomes. When a linear model is used at the post-secondary stage, the instrumental variable approach is used to control for endogeneity. When an order probit model is used at the post-secondary stage, the maximal likelihood approach is used. However, the maximal likelihood approach also relies on the exclusion assumption that is the essence of the instrumental variable approach.

3. Multi-stage education model

The multi-stage education model for estimation follows closely the well accepted theoretical specification. The human capital output at an earlier stage acts as an intermediate input in the education technology at a later stage and improves the learning efficiency thereof, i.e., there is the human capital dynamic linkage. More specifically, I assume the Markovian property that all relevant information of the entire schooling history is adequately summarized by the resulting education

---

[1] For exmaple, see Olsen and Farkas (1989), Jensen (1990), Carrasco (2001), Bingley and Walker (2001), and Kenkel and Terza (2001).

outcome at a given stage, i.e., the human capital stock is a state variable. Let the high school stage be denoted as stage 1, the education technology at high school is as follows:

$$h_{i1} = \mathbf{x}_i'\beta_1 + \mathbf{z}_i'\gamma + u_{i1} \tag{1}$$

where $\mathbf{x}_i$ is the vector of exogenous variables to control for observed heterogeneity, and $\mathbf{z}_i$ are the instruments measuring high school quality (to be discussed in detail later). The human capital output at high school is h_{i1}.

The central equation capturing the human capital dynamic linkage is that at the post-secondary stage, denoted as stage 2. Human capital output at high school acts as an intermediate input at the post-secondary stage and improves the learning efficiency. So students' human capital output from post-secondary schooling, h_{i2}, is given by the following technology:

$$h_{i2} = \alpha h_{i1} + \mathbf{x}_i'\beta_2 + u_{i2} \tag{2}$$

Here the parameter α captures the human capital dynamic linkage: How students' high school academic performance at high school affects their post-secondary schooling outcomes, controlling for the observed heterogeneity $\mathbf{x}_i$.

The unobserved factors at these stages are u_{i1} and u_{i1}. It is assumed that $E(u_{i1}|\mathbf{x}_i,\ \mathbf{z}_i) = E(u_{i2}|\mathbf{x}_i,\ \mathbf{z}_i) = 0$, $E(u_{i1}^2|\mathbf{x}_i,\ \mathbf{z}_i) = \sigma_1^2$, $E(u_{i2}^2|\mathbf{x}_i,\ \mathbf{z}_i) = \sigma_2^2$, and $E(u_{i1}u_{i2}|\mathbf{x}_i,\ \mathbf{z}_i) = \rho\sigma_1\sigma_2$. Notice that homoscedasticity of the random term within a given stage does not necessarily imply homoscedasticity across education stages, as uncertainty may play a bigger role in some stages than in others. If $\rho = 0$, we have $E(u_{i2}|h_{i1},\ \mathbf{x}_i) = 0$, there is no endogeneity problem, and equation (2) can be estimated directly without bias. If $\rho \neq 0$, there is endogeneity in h_{i1}, direct estimation of equation (2) is biased, and we have to utilize the instruments $\mathbf{z}_i$ in the two-stage least squares (2SLS) estimation of equations (1) and (2). For simplicity, it is further assumed that the random terms are uncorrelated across individuals, i.e., $E(u_{is}u_{jt}|\mathbf{x}_i,\ \mathbf{z}_i,\ \mathbf{x}_j,\ \mathbf{z}_j) = 0$ for $i \neq j$ and $s,\ t = 1,\ 2$.

3.1. The dataset

The data used for estimation comes from the National Educational Longitudinal Study of 1988 (hereinafter referred to as NELS88). In that study, a nationally representative sample of eighth-graders was first surveyed in the spring of 1988. Later, a sample of these respondents was resurveyed through four follow-ups in 1990, 1992, 1994, and 2000 respectively. This study covers the transitional process from high school to post-secondary education. The dataset provides a rich array of information by surveying not only the students, but also their parents, their teachers and the school administrators at the high school stage. It allows good control of observed heterogeneity.

Relevant data are extracted from the base year, the second follow-up and the fourth follow-up surveys. Table 1 defines the variables. Human capital output at the high school stage is measured by the standardized test score at the twelfth grade (TEST). In the dataset, there are test scores on four subjects: reading, mathematics, science and history/citizenship/geography. Standardized, the test scores at the four subjects have almost indistinguishable means and standard deviations. These test scores are highly correlated and preliminary principal component analysis indicates that only one (an arbitrary one) has eigenvalue greater than 1 (the remaining less than 0.1). Consequently TEST takes the average of the available standardized test scores on the four subjects for each student to prevent collinearity and to minimize missing data problems.

Human capital output at the post secondary stage is measured by the highest degree attained (PSEDG). It is a categorical variable ranging from no post-secondary education up to the highest Ph. D. and professional degrees. Later both a linear model and an ordered probit model is used to fit PSEDG at the post-secondary stage. The linear model facilitates the implementation of the instrumental variable approach. The ordered probit model generates a realistic range of predictions. Estimation results of both models share the same qualitative features.

Following the literature, the vector of exogenous variables $\mathbf{x}_i$ includes the student's gender (MALE), race/ethnicity (API, HISPANIC, BLACK, WHITE and RACEOTH), parental highest education levels (FAED and MOED), total family income (INC, averaged between year 1987 and year 1991), and the student's own expectation about his education achievement (EDEXP). These variables allow for good control of the observed heterogeneity.

3.2. Identification strategy

The vector of instruments $\mathbf{z}_i$ contains variables measuring the high school quality, namely whether the high school is public or private (PUBSCH), the location of the high school (URBAN, SUBURBAN and RURAL), and average class size measured as the composite student/teacher ratio (RATIO). High school quality has direct impact on high school but not post-secondary outcomes, hence these variables are excluded at the post-secondary stage. This exclusion restriction is the foundation of identification, both in the instrumental variable approach when a linear model is used at the post-secondary stage, and the maximal likelihood approach when an ordered probit model is used.

On the other hand, the validity of the instruments depends critically on their exogeneity, i.e., they should be uncorrelated with u_{i2}. Obviously, the quality of high school is a parents' choice. Parents may implicitly choose the high school quality by making their residential choices, or explicitly by choosing private schools. These choices may be based on factors unobservable to us but observed by parents, and these factors may have impacts both at the high school stage and the post-secondary stage. When this is the case, the exogeneity of the instruments is not guaranteed.

However, we can argue that the violation of exogeneity of the instruments may not be so severe a problem. Consider the three instruments one by one. First the location of high school (URBAN, SUBURBAN and RURAL) is a very crude categorization. Although common sense tells us that an average suburban high school is of better quality than an average inner city or rural high school, there is a wide spectrum of quality within each category nevertheless. So parental choice is more likely to be a finer decision within a category rather than choosing among the three categories directly. This reduces the possible endogeneity. Next the class size is measured as the school-wise student teacher ratio (RATIO), instead of the actual class size. The endogeneity of the class size arises when either a more able student is put in a smaller class (promotive assignment), or a less able student is put in a smaller class (remedial assignment). The school-wise class size is of course correlated but not perfectly correlated with the actual class size. The lack of perfect correlation mitigates the endogeneity problem. Last, when choosing between public and private high schools (PUBSCH), parents have more concerns than only the academic performance: religious reasons (roughly three quarter of the private schools have religious affiliations), athletics, loyalty to one's alma mater, etc. These reasons other than academic performance may hardly affect education outcomes at the post-secondary stages, so again the endogeneity problem is mitigated. I also conduct over-identificaiton tests with regard to the validity of the instruments. As in Woodbridge (2002), the test for validity of the over-identification restrictions is obtained as NR^2 from the OLS regression of the 2SLS residuals on all the exogenous variables (including all the instruments), where N is the sample size and R^2 is the usual R-squared. Under the null of exogenous instruments, it is asymptotically chi-square distributed, with the degree of freedom equal to the number of over-identification restrictions. All but one test fail to reject the null of valid over-identification restrictions (the remaining one has a p-value of 0.0786, significant at 10% but insignificant at 5% significance level).

3.3. Summary statistics

Out of total 12,144 observations in the dataset, 3,068 observations lack data on the human capital measurements at high school and post-secondary schooling (TEST and PSEDG); additional 900 observations have missing data on one or more of the exogenous variables in the vector $\mathbf{x}_i$, namely MALE, API, HISPANIC, BLACK, WHITE, RACEOTH, FAED, MOED, INC and EDEXP; yet another 98 observations miss information on the instrument RATIO. As discussed in Allison (2002), observations with missing dependent variables do not contribute to estimation efficiency, I drop the 3,068 observations from the sample. Also with the sufficiently big remaining sample size, I can afford the luxury of list-wise deletion of the additional 998 observations with missing data on the explanatory variables, in that list-wise deletion leads to consistent estimation if the data are missing at random. The actual sample used for estimation has 8,078 observations, or 66.5 % of the original dataset. The summary statistics are reported in table 2.

As can be seen, 47.4% of the students are male; 6.7% Asian/Pacific islander, 11.5% Hispanic, 8.4% African American, 69.9% white and 3.6 % American Indian or of multiple races. Most of the students' parents either just graduate from high school or have some post-secondary education but shy of a Bachelor's degree, yet most of the students themselves expect to finish college and get a Bachelor's degree. On average a family has annual income of 45.5 thousand dollars. In terms of high school quality, 81.6% of the students attend public schools; 24.3% of the high schools are in urban, 41.7% suburban and 34.0% rural area; and the average class size is 18 students. At the twelfth grade, the average student scored 51.8 in his standardized test; and twelve year after high school, most students either have some post-secondary education or have completed college with a Bachelor's degree.

4. IV estimation with a linear model

Although the human capital at the post-secondary stage (PSEDG) is a categorical variable, for the ease of implementation of the instrumental variable approach, a linear model is used to fit it in this section. The estimation of an ordered probit model is deferred to the next section. With a linear model, the instrumental variable approach is just the familiar two-stage least squares (2SLS). As robustness check, the whole sample is further divided into subsamples along two dimensions: gender and geographic location (the high school location is also the student's residential location). Estimation results are reported parallel for the full sample and the subsamples.

Table 3 reports the first stage results. Consistent with the existing literature, family background such as parental education levels and family income has significantly positive impact on students' high school outcomes. Even more prominent is the impact of students' own expectation on education attainment. In terms of schooling quality, private high schools outperform public high schools; schools with smaller class size outperform schools with bigger class size; and there is mixed and insignificant evidence on the effect of school location. These relationships remain stable across subsamples. Overall the exogenous variables explain considerable fraction of the total variation in TEST (R^2 ranges from 0.30 to 0.41).

Table 4 reports the second stage results using the instrumental variable approach, i.e., the fitted value of TEST from the first stage is used as an regressor. The estimated human capital dynamic linkage is quite substantial. In the full sample, a 1% increase in the high school test score leads to a 6.36% increase in the post-secondary outcome; more equivalently, it takes roughly 16% increase in the high school test score to move the post-secondary attainment one scale up. Also the human capital dynamic linkage remains as a stable relationship across subsamples, with its magnitude ranging from 4.57% in the female subsample up to 7.30% in the rural subsample. Of course one drawback of the linear model is that predicted PSEDG may fall outside the realistic range from 1 to 5, especially at the lower end where predicted values may be negative. That motivates us to turn to an ordered probit model in the next section. However, the estimation consistency is guaranteed as long as the instruments are valid.

Over-identification tests are conducted, and p-values are reported in Table 4. As discussed before, the over-identification test is obtained as NR^2 from the OLS regression of the 2SLS residuals on all the exogenous variables (including all the instruments). Under the null of exogenous instruments, it is asymptotically chi-square distributed, with the degree of freedom equal to the number of over-identification restrictions. A high p-value translates as fail to reject the null of valid over-identification restrictions. Five of the six tests have p-values above any practical significance level, and the remaining one (female subsample) has a p-value of 0.0786, significant at 10% but insignificant at 5% significance level. Overall we have reasonable confidence that our choice of instruments is not seriously flawed.

A by-product of the two-stage least squares is the correlation coefficient ρ, which is estimated from the residuals at the two stages. Surprisingly the error correlation between the high school stage and the post-secondary stage is highly negative. The negative error correlation may arise when students with better high school performance self select into more difficult (and probably also more rewarding) majors, while the variable PSEDG is not properly adjusted by the major-specific difficulty level. The negative error correlation may also arise when there is intertemporal allocation of efforts across stages. With a more or less fixed amount of total efforts, students who work hard in high school tend to slack relatively in post-secondary education, yet students who dabble in high school tend to focus more in post-secondary education. This may simply reflect the nature of a mean reversion process.

Another interesting result worth noting is that when the human capital dynamic linkage is directly included as an intermediate input, family background variables such as FAED, MOED, INC and EDEXP either lose their significance or turn negative at the post-secondary stage. Return to the example given in Section 2. Now it becomes clear that when students with better educated parents do better in college, the major effect does not come from parent-child interaction at the current stage; instead it comes from parent-child interaction at earlier stages. In short, the indirect effect is the dominant component in the long-term impact of family background. And the most effective education policy would be those targeting early intervention instead of later remedy.

As specification tests, Table 5 reports the OLS estimation results and Hausman tests for endogeneity of the regressor h_{i1}. The null is that $\rho = 0$, namely there is no endogeneity problem. Then OLS is both consistent and efficient under the null, but inconsistent when the null fails to hold; while IV is consistent regardless of the null. As can be expected, the point estimate of the human capital dynamic linkage α under OLS is substantially smaller than that under IV, ranging from 20.4% up to 39.1% in magnitude. On the other hand, the Hausman test rejects the null in three cases (the full sample, male subsample and urban subsample), and fails to reject the null in the remaining three cases (female subsample, suburban subsample and rural subsample). When the Hausman test fails to reject the null, it is mainly because the big difference in the covariance matrix under IV and OLS, which hints that the instruments may not be as strong in the latter three cases as in the former three.

In the next section, an ordered probit model is used at the post-secondary stage, and the maximum likelihood approach is used for estimation. Like IV, the MLE relies on the exclusion restriction for identification of the human capital dynamic linkage. But it has two advantages. First, with an ordered probit model, the predicted value of PSEDG is a probability and always falls within the realistic range, unlike that in the linear model. Second, MLE jointly estimates the system of equations (1) and (2), and utilizing the error correlation ρ may improve estimation efficiency. Also the next section serves as a robustness check of the human capital dynamic linkage with regard to functional form choices.

5. MLE estimation with an ordered probit model

To formulate the likelihood function, the error terms u_{i1} and u_{i2} in (1) and (2) are assumed jointly normally distributed, with means 0, variances σ_1^2 and 1 (σ_2 is a scale factor and cannot be independently estimated from α and β_2), and correlation coefficient ρ. The results of joint estimation of equations (1) and (2) are reported in Table 6.

Overall the MLE results share the same qualitative feature with the IV estimates. At the high school stage, family background has significantly positive impact on students' test scores; and students attending schools with higher quality (private schools and smaller class size) have better performance. At the post-secondary stage, the first result to notice is that the human capital dynamic linkage is again significantly positive, and this relationship is stable across subsamples. Next, when the human capital dynamic linkage is taken into account, family background loses its significance or turn negative at the post-secondary stage, again suggesting that the indirect effect channeled through the human capital dynamic linkage is the dominant component in the long-term effect of family background. Thirdly, under joint estimation, the error correlation ρ is significantly negative, strengthening our confidence built from the Hausman tests under IV that there is error correlation across education stages, and hence endogeneity has to be controlled for estimation consistency.

Table 7 translates the parameters in the ordered probit model into marginal effects. As can be seen, in the full sample, a 1% increase in the test score reduces the probability that the student has no post-secondary education by 2.03%, it also reduces the probability of some post-secondary education, but not a Bachelor's degree, by 0.75%; on the other hand, it increases the probability of getting a Bachelor's degree, a Master's degree, a Ph. D. or professional degree by 1.76%, 0.63% and 0.39% respectively. The marginal effects in the subsamples are similar. The results again suggest that better high school academic performance improves students' learning efficiency and hence leads to higher degrees at post-secondary education. There is strong human capital dynamic linkage from high school to post-secondary schooling.

6. Discussion and conclusion

This paper proposes a multi-stage education model to identify the human capital dynamic linkage from high school to post-secondary schooling. The parameter α has enormous policy importance. Since most education programs cover only a part of the entire schooling process, focusing on the direct effect, as has been done in the existing literature on program evaluation, overlooks a substantial part of the total benefit.

Consider an education program that is effective at the high school stage. Its total effect consists of the direct component and the indirect component. The direct effect arises within the high school stage, when policy intervention improves the human capital output by increasing certain education input, call it x, at this stage. The direct effect hence can be represented as $\frac{\partial h_{i1}}{\partial x_i}$. The indirect effect arises at the post-secondary stage, when better human capital output at high school improves the learning efficiency and education outcomes at the post-secondary stage. The indirect effect hence can be represented as $\frac{\partial h_{i2}}{\partial h_{i1}}\frac{\partial h_{i1}}{\partial x_i} = \alpha\frac{\partial h_{i1}}{\partial x_i}$.

The first implication of this is that, no matter how heavily discounted, the total effect (the sum of the direct and the indirect effect) is always bigger than the direct effect alone. Hence conventional program evaluation that focuses only on the direct effect gives underestimate of the total effect of any program. So when the program evaluation results are used to choose among programs that are effective at different education stages, say, between high school and post-secondary education, it is biased against the high school intervention program, because a bigger fraction of the total effect would be indirect and ignored in the program evaluation. However, this paper finds significantly positive α, suggesting strong indirect effect. So ranked by the total effect, it is more likely that programs effective at the high school stage have higher returns; yet ranked by the direct effect only, programs at the post-secondary stage may seem more beneficial. It is in this sense that conventional program evaluation may generate misleading policy recommendations, and the indirect effect channeled through the human capital dynamic linkage should be properly taken into account.

Next, when human capital is measured as a unidimensional variable, the human capital dynamic linkage α is a constant, and hence the indirect effect is of fixed proportion to the direct effect. This seems to suggest that if the conventional program evaluation results are used to choose programs effective at the same education stage, then higher direct effect also implies higher total effect, and the policy recommendation may be justified. However, in this study, although the lack of information prevents us from identifying differential human capital dynamic linkage within different subjects, there should be no a priori belief that the human capital dynamic linkage is the same across subjects. We can conjecture that when it does differ across subjects, again conventional program evaluation results would be biased against programs targeting subjects with strong human capital dynamic linkage. The policy recommendation may again be inefficient since it prefers maximal direct effect to maximal total effect.

This multi-education model provides a general vehicle for evaluating long-term effects of education programs. It assumes a stable relationship between the test score and true human capital level. Of course this stable relationship may no longer hold if under certain programs, individuals (school administrator, teachers or the students) have incentives to manipulate test scores. For example, if teachers get rewarded for good test scores of their students, they may focus on teaching exam-taking skills. When this happens, students do not acquire as much knowledge and skills that really contribute to their future learning efficiency, and the human capital dynamic linkage is weakened. More detailed analyses are needed to identify the relevant human capital dynamic linkage for evaluating the indirect effect of programs.

The paper is only a first step towards understanding the hierarchical structure in the education process. The identified parameter α of human capital dynamic linkage captures only the local relationship from high school to post-secondary schooling. When more stages are introduced before high school, our understanding may be enhanced. For example, in the current two-stage setup, the high school stage takes the reduced form specification and finds significantly positive impact of family background. If an additional stage is introduced before high school, then we may know better whether it is mainly direct impact or indirect impact, as we have discovered at the post-secondary stage. My guess would be that still the indirect effect dominates, with the major impact dating back to early childhood development. Nonetheless, more data and more careful studies are needed before we truly understand the hierarchical structure of the entire education system.

References

Allison, Paul D. "Missing Data. Series: Quantitative Applications in the Social Sciences". Sage Publications, 2001.

Angrist, Joshua and Lavy, Victor. "Using Maimonides' Rule to Estimate the Effect of Class size on Children's Academic Achievement", Quarterly Journal of Economics, May 1999, 114(2), pp. 533-575.

——— and ———. "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools". Journal of Labor Economics, Apr. 2001, 19(2), pp. 343-369.

Ben-Porath, Yoram. "The Production of Human Capital and the Life Cycle of Earnings". Journal of Political Economy, Aug., 1967, 75(4), pp. 352-365.

Bingley, Paul and Walker, Ian. "Household Unemployment and the Labor Supply of Married Women". Economica, May 2001, 68(270), pp. 157-185.

Cameron, Stephen V. and Heckman, James J. "The Nonequivalence of High School Equivalents". Journal of Labor Economics, Jan., 1993, 11(1), pp. 1-47.

——— and ———. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males". Journal of Political Economy, Apr. 1998, 106(2), pp. 262-333.

Card, David and Krueger, Alan B. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States". Journal of Political Economy, Feb. 1992, 107(1), pp. 151-200.

Carrasco, Raquel. "Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the Effect of Fertility on Female Labor Participation". Journal of Business and Economic Statistics, Oct. 2001, 19(4), pp. 385-394.

Currie, Janet and Thomas, Duncan. "Does Head Start Make a Difference?" American Economic Review, June 1995, 85(3), pp. 341-364.

——— and ———. "School Quality and the Longer-term Effects of Head Start". Journal of Human Resources, Autumn 2000, 35(4), pp. 755-774.

——— and ———. "Longer-term Effects of Head Start". American Economic

Review, Sep. 2002, 92(4), pp. 999-1012.

Hoxby, Caroline M. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation", Quarterly Journal of Economics, Nov. 2000, 115(4), pp. 1239-1285.

Jensen, Eric R. "An Econometric Analysis of the Old-Age Security Motive for Childbearing". International Economic Review, Nov. 1990, 31(4), pp. 953-968.

Kenkel, Donald S. and Terza, Joseph V. "The Effect of Physician Advice on Alcohol Consumption: Count Regression With an Endogenous Treatment Effect". Journal of Applied Econometrics, March-April 2001, 16(2), pp. 165-184.

Krueger, Alan B. "Experimental Estimates of Education Production Functions". Quarterly Journal of Economics, May 1999, 114(2), pp. 497-532.

Lucas, Robert E. Jr. “ On the Mechanics of Economic Development”, Journal of Monetary Economics, July 1988, 22(1), pp. 3-42.

Mare, Robert D. "Social Background and School Continuation Decisions". Journal of American Statistical Association, June 1980, 75(370), pp. 295-305.

Olsen, Randall J. and Farkas, George. " Endogenous Covariates in Duration Models and the Effect of Adolescent Childbirth on Schooling". Journal of Human Resources, Winter 1989, 24(1), pp. 39-53.

Todd, Petra E. and Wolpin, Kenneth I. "On the Specification and Estimation of the Production Function for Cognitive Achievement". Economic Journal, Feb. 2003, 113(485), pp. F3-F33.

Woodbridge, Jeffrey M. "Econometric Analysis of Cross Section and Panel Data". MIT press, 2002.

Table 1 Definition of the Variables

MALE	Gender of student, 1 male, 0 female
API	Whether student id Asian/Pacific islander, 1 yes, 0 no
HISPANIC	Whether student is Hispanic, 1 yes, 0 no
BLACK	Whether student is African American, 1 yes, 0 no
WHITE	Whether student is white, 1 yes, 0 no
RACEOTH	Whether student is American Indian or more than one race, 1 yes, 0 no
FAED	Father's highest education level, 1 not finish high school, 2 graduated high school, 3 some college – no degree, 4 Bachelor's degree, 5 Master's degree, 6 Doctorate or professional degree
MOED	Mother's highest education level, coded the same as FAED
INC	Total family income (thousand dollars), average of 1987 and 1991
EDEXP	Student's own expectation of education attainment, 1 won't finish high school, 2 will finish high school, 3 some college – no degree, 4 will finish college, 5 higher schooling after college
PUBSCH	Whether student's high school is public, 1 yes, 0 no
URBAN	Whether student's high school in central city, 1 yes, 0 no
SUBURBAN	Whether student's high school in suburban area, 1 yes, 0 no
RURAL	Whether student's high school in rural area, 1 yes, 0 no
RATIO	Composite student/teacher ratio in high school
TEST	Standardized test scores at the twelfth grade, average among reading, mathematics, science and history/citizenship/geography
PSEDG	Highest post-secondary degree attained in 2000, 1 no post-secondary education, 2 some post secondary education, 3 Bachelor's degree, 4 Mater's degree, 5 Doctorate or professional degree

Table 2 Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
MALE	0.4740	0.4994	0	1
API*	0.0670	0.2500	0	1
HISPANIC	0.1148	0.3187	0	1
BLACK	0.0837	0.2769	0	1
WHITE	0.6986	0.4589	0	1
RACEOTH	0.0360	0.1864	0	1
FAED	2.6587	1.4847	1	6
MOED	2.5384	1.2976	1	6
INC**	45.4950	36.3871	1	200
EDEXP	3.8732	0.9107	1	5
PUBSCH	0.8158	0.3877	0	1
URBAN*	0.2434	0.4291	0	1
SUBURBAN	0.4171	0.4931	0	1
RURAL	0.3396	0.4736	0	1
RATIO**	17.7014	4.7177	10	30
TEST**	51.7754	8.7773	29.35	69.78
PSEDG	2.2668	0.8041	1	5

Note: * Default group in regression; **Natural logarithm used in regression.

Table 3 High School Stage

dependent variable	Whole sample		MALE		FEMALE	
TEST	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
MALE	0.0231	7.14				
HISPANIC	-0.0741	-9.34	-0.0611	-5.36	-0.0835	-7.53
BLACK	-0.1285	-14.74	-0.1386	-10.59	-0.1190	-10.17
WHITE	-0.0186	-2.93	-0.0138	-1.52	-0.0214	-2.40
RACEOTH	-0.0689	-6.59	-0.0492	-3.24	-0.0846	-5.89
FAED	0.0143	9.66	0.0123	5.41	0.0159	8.17
MOED	0.0083	5.09	0.0080	3.31	0.0090	4.08
INC	0.0289	11.16	0.0247	6.00	0.0322	9.66
EDEXP	0.0597	28.76	0.0630	19.69	0.0568	20.92
PUBSCH	-0.0340	-7.77	-0.0389	-6.16	-0.0307	-5.09
SUBURBAN	0.0007	0.17	0.0100	1.60	-0.0077	-1.37
RURAL	-0.0019	-0.41	0.0031	0.44	-0.0070	-1.11
RATIO	-0.0267	-4.37	-0.0110	-1.22	-0.0423	-5.16
INTERCEPT	3.6667	153.72	3.6481	103.32	3.7112	115.54
R-Squared	0.3386		0.3175		0.3586	
# of Obs.	8078		3829		4249	

dependent variable	URBAN		SUBURBAN		RURAL	
TEST	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
MALE	0.0153	2.40	0.0296	5.88	0.0216	3.81
HISPANIC	-0.0778	-5.99	-0.0679	-5.59	-0.0705	-3.59
BLACK	-0.1177	-7.86	-0.1267	-9.35	-0.1392	-7.03
WHITE	-0.0062	-0.57	-0.0265	-3.00	-0.0164	-0.98
RACEOTH	-0.0787	-3.67	-0.0617	-3.91	-0.0701	-3.29
FAED	0.0136	4.67	0.0167	7.57	0.0112	4.02
MOED	0.0065	2.16	0.0060	2.41	0.0131	4.25
INC	0.0293	5.68	0.0330	8.22	0.0231	5.15
EDEXP	0.0562	13.02	0.0603	18.08	0.0617	18.00
PUBSCH	-0.0406	-5.61	-0.0274	-4.30	-0.0287	-2.24
RATIO	-0.0231	-2.04	-0.0314	-3.52	-0.0200	-1.54
INTERCEPT	3.6767	80.46	3.6593	104.84	3.6490	73.34
R-Squared	0.4052		0.3183		0.2963	
# of Obs.	1966		3369		2743	

Table 4 Post-Secondary Stage – Linear Model (IV)

dependent variable	Whole sample		MALE		FEMALE	
PSEDG	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
TEST	6.3581	8.42	6.6829	6.26	4.5696	5.91
MALE	-0.2501	-8.99				
HISPANIC	0.2222	3.00	0.2046	2.04	0.1072	1.24
BLACK	0.5470	5.07	0.6475	3.90	0.2897	2.68
WHITE	-0.1052	-2.54	-0.1207	-2.01	-0.1236	-2.50
RACEOTH	-0.0126	-0.15	-0.1000	-0.89	-0.0730	-0.75
FAED	-0.0140	-0.93	-0.0146	-0.69	0.0133	0.79
MOED	-0.0123	-0.98	-0.0048	-0.25	-0.0058	-0.42
INC	-0.0215	-0.73	0.0050	0.12	0.0140	0.43
EDEXP	-0.1237	-2.59	-0.1714	-2.38	0.0030	0.06
INTERCEPT	-21.9872	-8.24	-23.4414	-6.17	-15.6329	-5.72
ρ	-0.7443		-0.7897		-0.5295	
# of Obs.	8078		3829		4249	
OI test	p-value=0.4893		p-value=0.3278		p-value=0.0786	

dependent variable	URBAN		SUBURBAN		RURAL	
PSEDG	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
TEST	6.9989	5.70	6.1757	4.90	7.3036	2.57
MALE	-0.1988	-4.11	-0.2888	-5.84	-0.2598	-3.43
HISPANIC	0.2318	1.82	0.1325	1.16	0.5179	2.16
BLACK	0.5237	3.08	0.4691	2.70	0.9597	2.30
WHITE	-0.2228	-3.01	-0.0759	-1.21	0.1027	0.84
RACEOTH	0.1401	0.79	-0.0976	-0.82	0.2436	0.98
FAED	-0.0288	-1.07	-0.0210	-0.79	-0.0013	-0.03
MOED	-0.0259	-1.11	-0.0083	-0.48	-0.0211	-0.47
INC	-0.0718	-1.29	0.0105	0.20	-0.0402	-0.54
EDEXP	-0.1697	-2.23	-0.1268	-1.59	-0.1681	-0.94
INTERCEPT	-24.0183	-5.52	-21.3458	-4.81	-25.7003	-2.54
ρ	-0.7791		-0.7294		-0.8068	
# of Obs.	1966		3369		2743	
OI test	p-value=0.4425		p-value=0.3147		p-value=0.6005	

Table 5 Post-Secondary Stage – Linear Model (OLS and Hausman test)

dependent variable	Whole sample		MALE		FEMALE	
PSEDG	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
TEST	1.5707	32.10	1.3648	20.10	1.7890	25.67
MALE	-0.1383	-9.79				
HISPANIC	-0.1191	-3.38	-0.1099	-2.25	-0.1201	-2.37
BLACK	-0.0516	-1.37	-0.0782	-1.47	-0.0320	-0.60
WHITE	-0.1715	-5.85	-0.1713	-4.23	-0.1718	-4.06
RACEOTH	-0.3299	-7.59	-0.3556	-5.88	-0.3014	-4.88
FAED	0.0612	9.49	0.0585	6.30	0.0619	6.93
MOED	0.0305	4.36	0.0402	4.07	0.0208	2.12
INC	0.1349	12.19	0.1611	9.86	0.1121	7.48
EDEXP	0.1669	18.69	0.1697	13.06	0.1630	13.29
INTERCEPT	-5.0564	-28.58	-4.5034	-18.12	-5.7978	-23.08
Hausman test of endogeneity	p-value<0.0001		p-value=0.0028		p-value=0.1482	

dependent variable	URBAN		SUBURBAN		RURAL	
PSEDG	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
TEST	1.6529	16.15	1.5705	20.64	1.5000	18.42
MALE	-0.1178	-4.22	-0.1502	-6.79	-0.1347	-5.53
HISPANIC	-0.1615	-2.71	-0.1770	-3.39	0.1126	1.46
BLACK	-0.0920	-1.47	-0.0914	-1.56	0.1504	1.89
WHITE	-0.1899	-3.68	-0.1790	-4.34	0.0164	0.25
RACEOTH	-0.2483	-2.74	-0.3717	-5.59	-0.1610	-1.98
FAED	0.0535	4.25	0.0606	6.40	0.0660	5.31
MOED	0.0153	1.18	0.0212	2.02	0.0564	4.21
INC	0.1155	5.08	0.1779	9.92	0.0988	5.53
EDEXP	0.1381	7.30	0.1537	10.87	0.1918	13.01
INTERCEPT	-5.0819	-13.95	-5.1142	-18.53	-5.0418	-16.69
Hausman test of endogeneity	p-value=0.0352		p-value=0.1972		p-value=0.9363	

Table 6 Post-Secondary Stage – Ordered Probit Model (MLE)

	Whole sample		MALE		FEMALE	
HS stage	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
MALE	0.0231	7.15				
HISPANIC	-0.0743	-9.37	-0.0622	-5.48	-0.0841	-7.59
BLACK	-0.1282	-14.71	-0.1396	-10.68	-0.1187	-10.16
WHITE	-0.0177	-2.80	-0.0136	-1.51	-0.0209	-2.36
RACEOTH	-0.0682	-6.53	-0.0493	-3.25	-0.0843	-5.89
FAED	0.0144	9.71	0.0122	5.40	0.0160	8.23
MOED	0.0084	5.12	0.0080	3.29	0.0090	4.06
INC	0.0288	11.19	0.0250	6.09	0.0316	9.54
EDEXP	0.0596	28.78	0.0628	19.75	0.0566	20.91
PUBSCH	-0.0337	-9.35	-0.0365	-6.56	-0.0337	-6.69
SUBURBAN	-0.0006	-0.23	0.0017	0.43	-0.0052	-1.08
RURAL	-0.0049	-1.57	-0.0031	-0.74	-0.0082	-1.60
RATIO	-0.0218	-4.78	-0.0160	-2.91	-0.0344	-4.30
INTERCEPT	3.6542	178.74	3.6657	136.60	3.6925	116.82
PSE stage						
TEST	7.2211	58.32	6.9962	53.41	6.9791	19.81
MALE	-0.2843	-11.48				
HISPANIC	0.2658	4.29	0.2363	2.87	0.2027	2.01
BLACK	0.6333	9.36	0.7047	7.44	0.4773	4.44
WHITE	-0.1110	-2.37	-0.1108	-1.76	-0.1579	-2.19
RACEOTH	-0.0185	-0.22	-0.0914	-0.79	-0.0701	-0.53
FAED	-0.0188	-1.41	-0.0225	-1.23	0.0090	0.42
MOED	-0.0148	-1.17	-0.0090	-0.50	-0.0127	-0.70
INC	-0.0178	-0.67	-0.0006	-0.01	0.0159	0.38
EDEXP	-0.1309	-4.12	-0.1866	-4.39	-0.0055	-0.10
CUT1	26.7691	71.30	26.1258	65.74	26.2032	22.94
CUT2	27.8709	84.26	27.0741	72.81	27.6181	26.26
CUT3	29.0452	97.36	28.1098	74.39	29.0975	30.27
CUT4	29.6004	101.96	28.5328	72.84	29.8704	32.54
ρ	-0.7824	-22.98	-0.8423	-21.67	-0.6098	-8.23
$\sigma 1$	0.1434	124.47	0.1486	82.82	0.1380	95.02

	URBAN		SUBURBAN		RURAL	
HS stage	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
MALE	0.0156	2.47	0.0296	5.88	0.0216	3.82
HISPANIC	-0.0775	-5.98	-0.0685	-5.67	-0.0704	-3.59
BLACK	-0.1173	-7.84	-0.1267	-9.36	-0.1391	-7.04
WHITE	-0.0063	-0.58	-0.0265	-3.00	-0.0163	-0.97
RACEOTH	-0.0788	-3.68	-0.0616	-3.90	-0.0701	-3.29
FAED	0.0138	4.78	0.0168	7.64	0.0112	4.02
MOED	0.0065	2.17	0.0060	2.40	0.0132	4.27
INC	0.0295	5.75	0.0330	8.26	0.0231	5.16
EDEXP	0.0562	13.05	0.0602	18.11	0.0617	18.05
PUBSCH	-0.0413	-6.38	-0.0296	-5.89	-0.0300	-3.36
RATIO	-0.0150	-1.92	-0.0263	-3.73	-0.0172	-1.98
INTERCEPT	3.6520	97.48	3.6469	117.55	3.6424	90.66
PSE stage						
TEST	7.6647	42.27	7.1394	34.06	7.2622	42.63
MALE	-0.2153	-4.37	-0.3329	-8.60	-0.2626	-6.06
HISPANIC	0.2635	2.45	0.1674	1.71	0.5279	3.89
BLACK	0.5817	4.95	0.5592	5.13	0.9607	7.09
WHITE	-0.2395	-2.95	-0.0745	-1.08	0.1047	0.92
RACEOTH	0.1544	0.87	-0.1088	-0.86	0.2236	1.46
FAED	-0.0354	-1.44	-0.0263	-1.23	-0.0031	-0.13
MOED	-0.0276	-1.19	-0.0125	-0.69	-0.0200	-0.76
INC	-0.0794	-1.63	0.0216	0.46	-0.0292	-0.70
EDEXP	-0.1827	-3.49	-0.1401	-2.82	-0.1486	-2.35
CUT1	27.9685	46.73	26.5134	42.21	27.1788	56.85
CUT2	29.0639	52.20	27.6190	50.03	28.1542	64.18
CUT3	30.1421	54.67	28.8334	58.89	29.1814	63.72
CUT4	30.7065	54.82	29.3805	62.22	29.6410	60.88
ρ	-0.8236	-15.85	-0.7711	-13.66	-0.8298	-13.65
σ1	0.1384	61.58	0.1444	77.55	0.1451	75.02

Table 7 Marginal Effects at the Post-Secondary Stage

			Full sample		
	Pr(PSEDG=1)	Pr(PSEDG=2)	Pr(PSEDG=3)	Pr(PSEDG=4)	Pr(PSEDG=5)
TEST	-2.0310	-0.7499	1.7594	0.6279	0.3936
MALE	0.0804	0.0286	-0.0690	-0.0245	-0.0154
HISPANIC	-0.0683	-0.0361	0.0609	0.0256	0.0178
BLACK	-0.1385	-0.1100	0.1212	0.0690	0.0583
WHITE	0.0306	0.0123	-0.0268	-0.0099	-0.0063
RACEOTH	0.0052	0.0019	-0.0045	-0.0016	-0.0010
FAED	0.0053	0.0020	-0.0046	-0.0016	-0.0010
MOED	0.0042	0.0015	-0.0036	-0.0013	-0.0008
INC	0.0050	0.0018	-0.0043	-0.0015	-0.0010
EDEXP	0.0368	0.0136	-0.0319	-0.0114	-0.0071

			MALE		
	Pr(PSEDG=1)	Pr(PSEDG=2)	Pr(PSEDG=3)	Pr(PSEDG=4)	Pr(PSEDG=5)
TEST	-2.2274	-0.4589	1.5064	0.5606	0.6193
MALE	-0.0704	-0.0221	0.0477	0.0203	0.0245
HISPANIC	-0.1754	-0.1000	0.1072	0.0659	0.1023
BLACK	0.0347	0.0081	-0.0235	-0.0090	-0.0102
WHITE	0.0299	0.0047	-0.0201	-0.0071	-0.0075
RACEOTH	0.0072	0.0015	-0.0048	-0.0018	-0.0020
FAED	0.0029	0.0006	-0.0019	-0.0007	-0.0008
MOED	0.0002	0.0000	-0.0001	0.0000	-0.0001
INC	0.0594	0.0122	-0.0402	-0.0150	-0.0165
EDEXP	-2.2274	-0.4589	1.5064	0.5606	0.6193

			FEMALE		
	Pr(PSEDG=1)	Pr(PSEDG=2)	Pr(PSEDG=3)	Pr(PSEDG=4)	Pr(PSEDG=5)
TEST	-1.4570	-1.2225	2.0839	0.4819	0.1137
MALE	-0.0386	-0.0405	0.0593	0.0158	0.0040
HISPANIC	-0.0788	-0.1092	0.1309	0.0441	0.0130
BLACK	0.0318	0.0292	-0.0468	-0.0114	-0.0028
WHITE	0.0152	0.0115	-0.0210	-0.0046	-0.0011
RACEOTH	-0.0019	-0.0016	0.0027	0.0006	0.0001
FAED	0.0026	0.0022	-0.0038	-0.0009	-0.0002
MOED	-0.0033	-0.0028	0.0047	0.0011	0.0003
INC	0.0011	0.0010	-0.0016	-0.0004	-0.0001
EDEXP	-1.4570	-1.2225	2.0839	0.4819	0.1137

	URBAN				
	Pr(PSEDG=1)	Pr(PSEDG=2)	Pr(PSEDG=3)	Pr(PSEDG=4)	Pr(PSEDG=5)
TEST	-1.9286	-1.1013	1.5655	0.8349	0.6295
MALE	0.0545	0.0304	-0.0440	-0.0233	-0.0176
HISPANIC	-0.0610	-0.0436	0.0492	0.0304	0.0251
BLACK	-0.1162	-0.1118	0.0863	0.0713	0.0704
WHITE	0.0593	0.0353	-0.0480	-0.0264	-0.0203
RACEOTH	-0.0361	-0.0253	0.0293	0.0177	0.0144
FAED	0.0089	0.0051	-0.0072	-0.0039	-0.0029
MOED	0.0069	0.0040	-0.0056	-0.0030	-0.0023
INC	0.0200	0.0114	-0.0162	-0.0087	-0.0065
EDEXP	0.0460	0.0263	-0.0373	-0.0199	-0.0150

	SUBURBAN				
	Pr(PSEDG=1)	Pr(PSEDG=2)	Pr(PSEDG=3)	Pr(PSEDG=4)	Pr(PSEDG=5)
TEST	-1.8764	-0.9198	1.7367	0.6369	0.4227
MALE	0.0878	0.0418	-0.0804	-0.0295	-0.0198
HISPANIC	-0.0414	-0.0248	0.0390	0.0159	0.0113
BLACK	-0.1148	-0.1052	0.1061	0.0610	0.0530
WHITE	0.0193	0.0100	-0.0180	-0.0068	-0.0046
RACEOTH	0.0299	0.0122	-0.0271	-0.0092	-0.0058
FAED	0.0069	0.0034	-0.0064	-0.0024	-0.0016
MOED	0.0033	0.0016	-0.0030	-0.0011	-0.0007
INC	-0.0057	-0.0028	0.0052	0.0019	0.0013
EDEXP	0.0368	0.0181	-0.0341	-0.0125	-0.0083

	RURAL				
	Pr(PSEDG=1)	Pr(PSEDG=2)	Pr(PSEDG=3)	Pr(PSEDG=4)	Pr(PSEDG=5)
TEST	-2.4635	-0.2047	1.6306	0.5544	0.4832
MALE	0.0894	0.0065	-0.0587	-0.0198	-0.0173
HISPANIC	-0.1528	-0.0521	0.1027	0.0486	0.0536
BLACK	-0.2380	-0.1311	0.1430	0.0938	0.1322
WHITE	-0.0361	-0.0019	0.0236	0.0078	0.0066
RACEOTH	-0.0711	-0.0139	0.0482	0.0188	0.0180
FAED	0.0010	0.0001	-0.0007	-0.0002	-0.0002
MOED	0.0068	0.0006	-0.0045	-0.0015	-0.0013
INC	0.0099	0.0008	-0.0065	-0.0022	-0.0019
EDEXP	0.0504	0.0042	-0.0334	-0.0113	-0.0099