



SEMIPARAMETRIC REDUCED FORM ESTIMATION OF TUITION SUBSIDIES

Hidehiko Ichimura
Christopher Taber

THE INSTITUTE FOR FISCAL STUDIES
DEPARTMENT OF ECONOMICS, UCL
cemmap working paper CWP01/02

Semiparametric Reduced Form Estimation of Tuition Subsidies

Hidehiko Ichimura
Department of Economics
University College London
and
Institute for Fiscal Studies

Christopher Taber¹
Department of Economics
and
Institute for Policy Research
Northwestern University

December 31, 2001

1 Introduction

The goal of this paper is to use a semiparametric reduced form model to estimate the effects of various tuition subsidies. This approach expands on the tuition subsidy example in Ichimura and Taber (2000) in a number of dimensions.

It has become common practice in the empirical literature to refer to any nonstructural empirical analysis as “reduced form.” This is not the traditional sense of the phrase. A classic reduced form analysis (see e.g. Marschak, 1953) first specifies a structural model and then derives the reduced form parameters in terms of the structural parameters. While many recent studies have asserted to taking a reduced form approach, the structural model which the reduced form model should correspond is rarely specified. We explicitly specify a structural model and use the implied reduced form structure to estimate the effect of tuition subsidy policies. Specifying the underlying model has the advantage of being explicit about the assumptions that justify the analysis. This avoids Rosenzweig and Wolpin’s (2000) criticism of work on natural ‘natural experiments’ that often leaves these conditions implicit. Our structural model is based on the model studied by Keane and Wolpin (1999). It is highly nonlinear and allows for more unobserved heterogeneity than the typical simultaneous equations framework that most previous work has used in reduced form estimation. Using the specified structural model, we examine the assumptions discussed in Ichimura and Taber (2000) to justify reduced form estimation of the policy effects.

2 Model

The model we pose is strongly based on Keane and Wolpin (1999). Hereafter we denote that paper by KW. The decision problem in the model begins at age 14 when students have completed eighth grade. We let each period represent one year so that $t = 0$ when the individual is 14. During that year the student chooses whether to attend college or not and how much to work. Let s_{it} be an indicator of whether individual i attends school in year t , ℓ_{it} the number of weeks worked in a year, and h_{it} the number of hours worked per week. Individuals have preference for consumption, work, and schooling so that the utility function is $u_{it}(c_{it}, s_{it}, S_{it}, \ell_{it}, h_{it}, Z_{it}^u)$, where S_{it} is cumulative years of school (i.e. $S_{it} = \sum_{\tau=0}^{t-1} s_{i\tau}$) and Z_{it}^u is an exogenous vector. For example KW let marriage and parental co-residence be elements of Z_{it}^u .

We assume that hourly wages take a standard Mincer form so that

$$\log(w_{it}) = \beta_0 + Z_{it}^w \beta_1 + \beta_2 L_{it} + \beta_3 L_{it}^2 + \alpha_i(S_{it}) + v_{it}, \quad (1)$$

where Z_{it}^w is a set of observable non time varying variables that influence wages, v_{it} is an unobservable variable, and L_{it} is accumulated weeks of experience as of the beginning of the year. That is $L_{it} = \sum_{\tau=0}^{t-1} \ell_{i\tau}$.

A key feature of this model is that we allow the schooling effect function $\alpha_i(\cdot)$ to arbitrarily vary across individuals in the model.

Accumulation of assets (K_{it}) by students is complicated by two features of the model. First students are borrowing constrained in two different sense. There is a limitation to the amount of debt that they can accumulate which may be individual specific. Let this amount be $-\bar{K}_i$. They also borrow at rate R_i^b but lend at rate R_i^ℓ . Second they may receive transfers from their parents. These transfers are individual specific and potentially depend on their schooling choices, college costs, their asset levels, and their wage rate. Denote this function as $P_{it}(s_{it}, S_{it}, \tau_{ic}, K_{it}, w_{it})$. We assume that the student takes this transfer function as given. This yields the budget constraint

$$K_{it+1} = K_{it} \left[R_i^b \mathbf{1}(K_{it} \leq 0) + R_i^\ell \mathbf{1}(K_{it} > 0) \right] - c_{it} \quad (2)$$

$$+ \ell_{it} h_{it} w_{it} + P_{it}(s_{it}, S_{it}, \tau_{ic}, K_{it}, w_{it}) - (\tau_{ic} - \pi_i) s_{it} \mathbf{1}(4 \leq S_{it} < 8) \quad (3)$$

$$K_{it+1} \geq -\bar{K}_i \quad (4)$$

where τ_{ic} is the monetary cost of attending college and π_i is the value of a tuition subsidy offered to individual i if they attend college. Note that the tuition subsidies we consider operate by changing this variable only. That is we do not allow for general equilibrium effects of the type described by Heckman, Lochner, and Taber (1998). This is valid if the simulations we consider relate to small programs that would only affect a small fraction of the labor market. We assume that there are no other monetary costs of college attendance.

The state variables in this model are assets, schooling, and experience (K_{it}, S_{it}, L_{it}). The control variables are consumption, school enrollment, weeks of labor, and hours of labor ($c_{it}, s_{it}, \ell_{it}, h_{it}$). Let $E_{it}(\cdot)$ represent the expectation of agent i with respect to information he has at age t . Thus one can write this model using Bellman's equation as

$$V_{it}(K_{it}, S_{it}, L_{it}) = \max_{c_{it}, s_{it}, \ell_{it}, h_{it}} \{u_{it}(c_{it}, s_{it}, S_{it}, \ell_{it}, h_{it}) + \delta E_{it}(V_{it+1}(K_{it+1}, S_{it+1}, L_{it+1}))\} \quad (5)$$

subject to (2), and (4). (6)

Individuals vary in the population because of heterogeneity in preferences (u_{it}), information (E_{it}), wages (w_{it}), parental transfers (P_{it}), and borrowing rates and asset level (R_i^b, K_{it}) and tuition level (τ_{ic}).

Loosely following Keane and Wolpin (1999) we assume that

$$\begin{aligned} u_{it}(c_{it}, s_{it}, S_{it}, \ell_{it}, h_{it}) &= \frac{c_{it}^\gamma}{\gamma} + \lambda_{is_1} s_{it} \mathbf{1}(S_{it} < 4) + \lambda_{is_2} s_{it} \mathbf{1}(4 \leq S_{it} < 8) + \lambda_{is_3} s_{it} \mathbf{1}(8 \leq S_{it}) \\ &\quad - \lambda_{i\ell} \frac{\ell_{it}^\alpha}{\alpha} - \lambda_{ih} \frac{h_{it}^\rho}{\rho} \\ P_{it}(s_{it}, S_{it}, \tau_{ic}) &= \nu_{i1} + \nu_{i2} \tau_{ic} s_{it} \mathbf{1}(4 \leq S_{it} < 8). \end{aligned}$$

Given this setup, we can solve for the optimal profile of schooling, consumption, savings, and leisure.

There are a number of sources of heterogeneity in the model. We model them in the following manner,

$$\begin{aligned} \lambda_{is_1} &= Z_i^{s'} \gamma^s + \theta_i^{s_1} + v_{it}^{s_1} \\ \lambda_{is_2} &= Z_i^{s'} \gamma^s + \theta_i^{s_2} + v_{it}^{s_2} \\ \lambda_{is_3} &= Z_i^{s'} \gamma^s + \theta_i^{s_3} + v_{it}^{s_3} \\ \lambda_{i\ell} &= Z_i^{\ell'} \gamma^\ell + \theta_i^\ell + v_{it}^\ell \\ \lambda_{ih} &= Z_i^{h'} \gamma^h + \theta_i^h + v_{it}^h \\ v_{i1} &= Z_i^{p'} \gamma^p + \theta_i^p + v_{it}^p \\ \tau_{ic} &= Z_i^{\tau'} \gamma^\tau + \theta_i^\tau + v_{it}^\tau \\ \alpha_i(S_{it}) &= \theta_{i0}^w + \theta_{i1}^w S_{it} + \theta_{i2}^w S_{it} \mathbf{1}(4 \leq S_{it}) + \theta_{i2}^w S_{it} \mathbf{1}(8 \leq S_{it}) \end{aligned}$$

where we assume that the error terms (terms with θ or v) are all independent with the observable variables (terms with Z). Let Z_i comprise all of the components of $(Z_i^w, Z_i^s, Z_i^\ell, Z_i^h, Z_i^p, Z_i^\tau)$.

3 Differences from the KW model

The main difference between our model and theirs is that our model does not require specifying the distribution of unobserved variables (i.e. the v_{it} and θ_i terms) and the agent's process of learning about the value of these unobservables. KW estimates the model by maximum likelihood which calls for specifying the joint distribution. Consistency of the estimator depends on the correct specification of this distribution.

Taber (2000) considers semiparametric identification of a similar class of dynamic programming schooling models and shows that the conditions are very strong and can not be met with the data at hand.

Instead we will exploit the variation in tuition variable and the assumption that (1) the variable is independent from the unobserved variable conditional on observed exogenous variables of the model, Z_i and (2) the variable is excluded from the outcome variables to estimate the impact of tuition subsidy policy.

If we could consistently estimate all the structural parameters, then we would be able to evaluate any policy that can be addressed within the model. We are not able to carry out this more general exercise. However, under the maintained assumptions (1) and (2) we will be able to consistently estimate the tuition subsidy policy effect without making distributional assumptions about the error terms.

4 Estimation

Note that in order to evaluate the tuition subsidy policy, we will not need to estimate the full structural parameters. We will estimate a semiparametric reduced form version of this model.

This is essentially a 6 index model. That is conditional on Z_i and age (A_{it}), for any potential outcome Y_{it} we can write

$$E(Y_{it}|Z_i, A_{it}, \pi_i) = F_Y \left(Z_i^{\tau'} \gamma^\tau - \pi_i, Z_i^{\ell'} \gamma^\ell, Z_i^h \gamma^h, Z_i^p \gamma^p, Z_i^s \gamma^s, Z_i^w \beta_1, A_{it} \right).$$

We will examine the following five outcome variables: college attendance, years of schooling, weeks worked per year, hours worked per week, and log-wage. The data is generated in the absence of a tuition subsidy ($\pi_i = 0$). Our goal is to simulate the effect of a tuition subsidy which can be summarized by π_i' which may vary across individuals. Our goal is thus to estimate $E(Y_{it}|Z_i, A_{it}, \pi_i')$.

We define the net tuition index $Z_i^{\tau'} \gamma^\tau - \pi_i$ to be

$$\text{tuition}_{17i} + \gamma^\tau \text{anycol}_{17i} - \pi_i$$

where tuition_{i17} is the average tuition paid at 2 year colleges in the state in which individual i lives at age 17 and anycol_{17i} is a dummy variable indicating whether there was a college in the county in which the individual lived when he/she was 17. We follow Cameron and Taber (2001) by assuming that this represents a financial cost of college as one can live at home

and attend a college if it is close. The structural model imposes two important restriction of that are being used here. The first is that these two variables in the cost of college index do not appear in any other indices. The second is that the tuition subsidy enters the index as specified above.

As formally discussed in the appendix, with an additional regularity condition, one can identify γ^τ and other indices when we normalize the unknown functions' first partial derivatives suitably. We do not need to impose any other exclusion restriction as it is not needed. Thus all variables in different indices except for the tuition index are the same:

$$Z_i^s = Z_i^\ell = Z_i^h = Z_i^p = Z_i^w = X_i.$$

For clarity we will write variables in the tuition index by T_i . Thus $Z_i^\tau = T_i$.

One strategy would be to estimate F_Y and the indices of the model up to normalizations as discussed in the appendix. This involves high dimensional nonlinear optimization. We found another strategy which reduces the dimensionality of the nonlinear optimization to one in our application. We discuss the precise relationship between the two approaches in the appendix.

Notice that for any outcome Y we can write the model even more flexibly as

$$E(Y_{it}|Z_i, A_{it}, \pi_i) = G_Y (T_i' \gamma^\tau - \pi_i, X_i, A_{it}).$$

This representation of the model is a classic semiparametric reduced form of the original structural model. The goal of this paper is to estimate the effect of tuition subsidy. We can identify/estimate the effect of the tuition subsidy on outcome Y in the following manner:

1. Using any particular outcome identify/estimate γ^τ and G_Y for that outcome.
2. Identify/Estimate G_Y corresponding to all the rest of outcomes.
3. Simulate $E(Y_{it}|Z_i, A_{it}, \pi_i') = G_Y (T_i' \gamma^\tau - \pi_i', X_i, A_{it})$.

The basic procedure for the first stage estimation of the index is similar to semiparametric least squares (Ichimura, 1993). Let \hat{G}_Y represent fit of a type of nonparametric regression of Y onto $(T' \gamma^\tau, X, A)$. Note that the γ^τ parameter is over identified as there are multiple outcomes. Simple models suggest that an easy way to estimate the effects of tuition on outcomes without sacrificing efficiency too much may be to use college attendance (D_i). Thus we estimate¹

$$\hat{\gamma}^\tau = \arg \min \sum_{i=1}^N \left(D_i - \hat{G}_D (\text{tuition}_{17i} + \gamma^\tau \text{anycol}_{17i}, X_i) \right)^2.$$

Age is dropped as D_i is constructed for individuals over 25 and under our assumption individual do not attend college after this.

In estimation we use a second order polynomial (including all interactions) as our flexible functional form for G_D . There are of course other possibilities. Note that given γ^r , use of quadratic function for \hat{G}_D implies that the rest of the parameters can be explicitly computed.

Stage 2 of the estimation is also performed using second order polynomial approximations.

One complication arises in interpretation of results. Wages at a given age depend on schooling both through its effect on experience and its direct effect. It is standard in the empirical literature to summarize the effect of schooling on wages holding experience constant. If we do not condition on experience, the effect of tuition subsidy on wages will depend on the level of experience and thus the age. However, given the structure above, this is not the case if we condition on experience. Since experience is endogenous (from the perspective of the model) conditioning on it is not trivial. However, given the form of the wage equation (1) above, there is a simple solution. Note that schooling only enters the model through $\alpha_i(S_i)$, which is fixed once individuals have entered the labor market. Thus one can deal with the endogeneity of S_i using a fixed effects strategy. There is another potential problem that labor supply and thus L_{it} will be correlated with v_{it} . We thus estimate β_2 and β_3 using fixed effects and instrumental variables with age and age squared as instruments for actual experience and experience squared. We can then report the effect of tuition subsidies on wages holding experience constant by using $\log(w_{it}) - \hat{\beta}_2 L_{it} - \hat{\beta}_3 L_{it}^2$ as our outcome variable. We then proceed with this outcome variable in exactly the same way that we do with the other outcomes.

5 Data

We use white, black, and Hispanic males from the 1979-1994 waves of the NLSY (National Longitudinal Survey of Youth). We exclude from the analysis the military subsample and the non-Black non-Hispanic disadvantaged sample. We use only male samples because their schooling decisions and labor supply are less complicated by fertility and labor market participation considerations.

Schooling observations begin at age 15 and extend at least through age 29 for all individuals included in our sample. Because information before 1978 is retrospective and limited, we confine our extract to males between ages 13 and 17 in January of 1978 in order to have reliable information on county of residence which is used to construct tuition and college

proximity.

The key variable in our analysis is tuition. A panel of annual records on tuition, enrollment, and location of all public two-and four-year colleges (including Universities) in the United States were constructed from the Department of Education's annual HEGIS and IPEDS "Institutional Characteristics" surveys. By matching location with a person's county of residence, we were able to determine the presence of a two-year college, four-year college, or either in the county of residence. A number of specialty colleges, generally with small enrollments less than 100, were excluded. Federal institutions, such as the Naval Academy, were also excluded. These variables can be measured at any age. For our analysis, however, they are measured at age 17 to avoid the obvious problem that people who attend college generally reside in the same county the college is located.

The other variables we control for in the models are minority status (i.e. black and Hispanic which are combined), family income, local income in the county in which the individual lived at age 17, AFQT score, fathers education, mothers education, and number of siblings. We use the same data as in Cameron and Taber (2001). It is discussed more extensively there.

6 Empirical Results

We present the results using the methodology and data described above. We consider three different types of tuition subsidies. The first is a straight \$1000 tuition subsidy. That is, every individual would receive a subsidy of \$1000 if he attends college. These results are presented in the first column of the table. The second is means tested and targeted towards low income families. Students whose parents income are less than \$20000 only are eligible for the \$1000 subsidy. These results are presented in column 2 and only averages for the eligible population (i.e. students from families who earn less than \$20000) are presented. The third subsidy is more complicated. Students from families who earn less than \$20000 are eligible for a \$2000 subsidy. The subsidy then phases out at a rate of \$100 per \$1000 of family income until it disappears at family income level \$40000. We present these results in column 3 for students from families that earn less than \$40000 only.

The results are presented in the Table. We examine several different outcomes. There are a couple aspects of the results to note. First notice that the direct effect of tuition subsidies on college attendance is smaller than what one often finds in straight regression (see Leslie and Brinkman, 1988, for literature review). This smaller effect results from combining

tuition with college attendance in the index, from the interactions with different variables, and from the extensive set of control variables. Interestingly the effects on hours per week is consistently negative. To give some manner of judging the magnitude of the wage effect we present the “implicit return to schooling effect” in the last row. This is simply the wage effect (controlling for experience) divided by the increase in years of schooling. This has the interpretation as the average return to a year of schooling averaged over the additional years of schooling that result from the subsidy. These are well within the range of the literature. Interestingly the effect on college enrollment is larger for the targeted subsidy than for the straight subsidy, but the return to schooling is very close. There are of course many other outcomes and cuts of the data one could construct using this methodology.

7 Conclusions

In this paper we perform semiparametric reduced form analysis to simulate the effects of several tuition subsidy programs on a variety of outcomes. This is classic reduced form estimation in that we present a full structural model but only estimate a reduced form of the model. Presentation of the full model is necessary to justify the use of the reduced form for policy simulation. This is an example of the type of direct estimation of policy effects described more generally in Ichimura and Taber (2000). The basic idea is to specify the minimal level of conditions necessary for identification of the policy effect and then use those conditions to simulate the effect. When one has the limited objective of simulating a particular policy this will often have the advantage over estimation of the full structural model that it requires weaker conditions. It may also be computationally easier to estimate the reduced form model than the full structural model. That is certainly the case in this example.

Using this approach we simulate several different tuition subsidies. As in other studies we find that tuition subsidies have a substantial effect on college attendance. This effect is likely to be larger when the subsidy is geared toward students from low income families. The effects of the subsidies on labor supply and wages is also documented.

References

- [1] Cameron, S., and Taber, C., (2001) “Estimating Borrowing Constraints using the Returns to Schooling,” unpublished manuscript.
- [2] Heckman, J., Lochner, L., and Taber, C. (1998) “General Equilibrium Treatment Effects,” *American Economic Review Papers and Proceedings*, 88, 381-386.
- [3] Ichimura, H. (1993), “Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models,” *Journal of Econometrics*, 58, 71-120.
- [4] Ichimura, H., and L. Lee (1991): “Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation,” in *Nonparametric and Semiparametric Methods in Economics and Statistics*, ed. by W.A. Barnett, J. Powell, and G. Tauchen (Cambridge, England: Cambridge University Press) 3-49.
- [5] Ichimura, H. and C. Taber (2000) “ Direct Estimation of Policy Impacts” NBER Technical Working Paper 254.
- [6] Leslie, L. and P. Brinkman (1988), *The Economic Value of Higher Education*, New York: Macmillan.
- [7] Marschak, J. (1953) “Economic Measurements For Policy and Prediction,” in *Studies in Econometric Method* edited by W. Hood and T. Koopmans, John Wiley 1–26.
- [8] Rosenzweig, M. and K. Wolpin (2000) “Natural ‘Natural Experiments’ in Economics”, *Journal of Economic Literature* 38(4): 827-874, Section 4 (p 853-862).

Appendix

Notice that the scale of coefficients are not identified and furthermore unless some exclusion restrictions are satisfied among indices there are further identification problems as discussed by Ichimura and Lee (1991). For our purpose, the crucial index is $Z_i^T \gamma^T$ and we assume that tuition variables affect decision only through this index. Under this assumption, the coefficients of this index is identified and thus the tuition policy effect is identified.

Let the model be

$$E(Y|X, T) = f(XB, T\theta)$$

where Y is a $r \times 1$ vector, X is a $1 \times k_1$ regressor vector, B is a $k_1 \times r$ matrix ($k_1 > r$), T is a $1 \times k_2$ regressor vector, and θ is a $k_2 \times 1$ vector ($k_2 > 1$), $f = (f^1, \dots, f^r)'$ and f^j ($j = 1, \dots, r$) is a continuously differentiable function from R^{r+1} into R . We assume that there is no common regressor among X and T and there is no a priori zero restriction in B matrix. We denote the partial derivative of f^j with respect to the p th argument by f_p^j .

In our application, $T_i^T \gamma^T$ is $T\theta$, the remaining 5 indices correspond to XB , and $r = 5$. Five outcomes are, whether one goes to college or not, and years of education, weeks worked, hours worked and log-wage.

We show the following: Let $\delta_{jp} = 1$ if $j = p$ and $\delta_{jp} = 0$ otherwise and denote the matrix with the j -pth element being $f_p^j(0)$ by A .

Theorem 1 *Let (X, T) be continuously distributed with its support being $R^{k_1+k_2}$. For any B , θ , and f , with θ_1 (the first element of θ) not zero, there exists a unique B^* , θ^* with $\theta_1^* = 1$, and \tilde{f} such that for $j = 1, \dots, r$ and $p = 1, \dots, r$*

$$\tilde{f}_p^j(0) = \delta_{jp}$$

and $f(XB, T\theta) = \tilde{f}(XB^, T\theta^*)$ almost everywhere if matrix A is nonsingular and B has full column rank.*

This normalization is chosen so that it can be imposed easily when we use polynomial expansion to approximate the unknown function f .

In estimation we use second order polynomial for the unknown functions. This would suggest using

$$\begin{aligned} D_i &= \delta_0 + \delta_1 \cdot (T_i^T \gamma^T) + X_i^T \gamma^p + \text{squared and cross terms of 6 indices} + e_{Di} \\ S_{it} &= \sigma_0 + \sigma_1 \cdot (T_i^T \gamma^T) + X_i^T \gamma^s + \text{squared and cross terms of 6 indices} + e_{si} \end{aligned}$$

$$\begin{aligned}
\ell_{it} &= \lambda_0 + \lambda_1 \cdot (T'_i \gamma^\tau) + \lambda_2 \cdot A_{it} \\
&\quad + X'_i \gamma^\ell + \text{squared and cross terms of 6 indices and } A_{it} + e_{\ell i} \\
h_{it} &= \eta_0 + \eta_1 \cdot (T'_i \gamma^\tau) + \eta_2 \cdot A_{it} \\
&\quad + X'_i \gamma^h + \text{squared and cross terms of 6 indices and } A_{it} + e_{hi} \\
\log w_{it} &= \omega_0 + \omega_1 \cdot (T'_i \gamma^\tau) + \omega_2 \cdot A_{it} \\
&\quad + X'_i \gamma^w + \text{squared and cross terms of 6 indices and } A_{it} + e_{wi}.
\end{aligned}$$

That there is only one linear term related with X_i and that the coefficient is set to 1 is the normalization the identification theorem suggests.

Note that the squared and cross terms of 6 indices which appear in each equation consists of cross terms and squared terms of the arguments in X_i and the cross terms of the arguments in X_i with $(T'_i \gamma^\tau)$ for D_i and S_i equations and analogous terms including age variable for the rest of the equations.

In principle, this model can be estimated by nonlinear least squares as we discussed in text. But implementing it will involve high dimensional nonlinear optimization problem. Hence we take an alternative approach.

Note that the nonlinearity mainly arises because we impose the model restriction that the coefficients in the ‘squared and cross terms’ are related to the linear terms. If we relax this restriction, the dimension of nonlinearity reduces to one.

The implication of this identification result is that the coefficients on the linear terms of X_i can be interpreted as the relevant index coefficients under the normalization specified in the theorem.

This observation can be used to (1) construct specification of the index restriction. Also, the estimation result can be used as the first step to impose the restriction implied by the index model.

Proof:

Let s be an $1 \times r$ vector t be a scalar and F be a nonsingular $r \times r$ matrix. Define function \tilde{f} on R^{r+1} into R^r by

$$\tilde{f}(s, t) = f(sF^{-1}, t \cdot \theta_1).$$

Let

$$F^{-1} = (F^1, \dots, F^r).$$

Then

$$\tilde{f}(s, t) = f(sF^1, \dots, sF^r, t \cdot \theta_1)$$

and it is easy to see that, denoting the p th ($p = 1, \dots, r$) element of F^j vector to be F_p^j for each $j = 1, \dots, r$,

$$\tilde{f}_p^j(0) = \sum_{s=1}^r f_p^j(0) F_p^s.$$

we have

$$AF^{-1} = I.$$

So $F = A$ uniquely solves the equation. This shows that the restrictions on the partial derivatives can be imposed without a loss of generality and that within the class of function \tilde{f} of the form denoted by $f(sF^{-1}, t \cdot \theta_1)$, \tilde{f} is uniquely determined.

Suppose $f(XB, T\theta) = g(XB^*, T\theta^*)$ holds almost surely and where g satisfies the partial derivative restrictions. Clearly B^* must be full rank since B is so from now on, we assume B^* has full column rank without a loss in generality.

By applying the single index identification argument conditional on X , θ^* must be proportional to θ so that there is a nonzero c such that $\theta^* = c \cdot \theta$. Since $\theta_1^* = 1$, $c = 1/\theta_1$.

We will show that $g(s, t) = f(sF^{-1}, t \cdot \theta_1) = \tilde{f}(s, t)$ almost surely for some nonsingular F . This, together with the earlier result shows the uniqueness.

Let

$$s = XB \text{ and } t = T\theta.$$

Since B has full column rank, there is some matrix that augments matrix B in such a way that the augmented matrix is nonsingular. Let (B, \tilde{B}) be such an augmented nonsingular matrix. Then $X(B, \tilde{B}) = (s, X\tilde{B})$ so that

$$XB^* = (s, X\tilde{B}) (B, \tilde{B})^{-1} B^*.$$

Then we have, denoting $u = X\tilde{B}$,

$$f(s, t) = g\left((s, u) (B, \tilde{B})^{-1} B^*, t/\theta_1\right)$$

almost surely. Taking derivative with respect to u and evaluating at the origin we have,

$$0 = AC'$$

or $C = 0$ where C is the bottom $k_1 - r$ rows of $(B, \tilde{B})^{-1} B^*$. This implies that B^* is a linear combinations of columns of B so that there is a matrix F such that $B^* = BF$. When B^* is nonsingular, F must be nonsingular so that we can write

$$f(XB^*F^{-1}, T\theta^* \cdot \theta_1) = g(XB^*, T\theta^*)$$

or $g(s, t) = f(sF^{-1}, t \cdot \theta_1)$ holds almost surely.

Simulated Effects of Tuition Subsidy
On Various Outcomes
Estimated from NLSY
Men Only

Outcome	\$1000 Subsidy ¹	Means Tests Subsidy ²	Phased Out Tuition Subsidy ²
College Attendance	0.0091	0.0323	0.0118
Years of Schooling	0.0254	0.0901	0.0326
Weeks Worked per Year	0.0175	0.0910	0.0545
Hours per Week	-0.0836	-0.4011	-0.2179
Log Wages (unconditional)	0.0021	0.0080	0.0029
Log Wages (cond. Exp.) ⁴	0.0025	0.0081	0.0017
Implicit Ret. to Schooling ⁵	0.0993	0.0904	0.0512
Fraction Eligible	100%	31%	69%

¹A \$1000 tuition subsidy is given to anyone attending college.

²A \$1000 tuition subsidy is given to individuals attending college whose family income is less than \$20000. The number reported is for the eligible population (family income <\$20000).

³The tuition is \$2000 for students whose family earns less than \$20000. The subsidy fades linearly to zero at \$40000. The numbers reported are for the eligible population (family income <\$20000).

⁴This number reports the effect of the policy on wages holding experience constant.

⁵The implicit return to schooling is simply the effect on log wages (conditional on experience divided by the effect on years of schooling).