Online Appendix: Government Expenditure on the Public Education System

Chao Fu[a], Shoya Ishimaru[b] and John Kennan[a*]

October 22, 2019

1 Data Details

All dollar values are measured in 2002 dollars.

1.1 Household Data

1.1.1 ELS 2002

We determine the household type x based on the status in the 2002 interview. Household income x_1 is categorized into 5 groups: \$20,000 or less, \$20,001-\$35,000, \$35,001-\$50,000, \$50,001-\$75,000, and \$75,001 or more. In the model, each income category is assigned its median income level in our sample from the 2000 Census, as summarized in Table B1. The household education level x_2 is an indicator for the presence of at least one adult with some college education. Minority status x_3 is an indicator signifying that a student is not White or Asian.

[*a:]University of Wisconsin and NBER.[b]University of Wisconsin. Contact: cfu@ssc.wisc.edu

Table B1: Discrete Income Levels

x_1	Income Range ($)	Median Income ($)	% in ELS sample
1	0-20,000	10,970	15.0
2	20,001-35,000	27,580	18.9
3	35,001-50,000	41,789	19.6
4	50,001-75,000	60,917	20.7
5	75,000-	103,636	25.8

We measure K-12 achievement k_1 based on the standardized math test test score and high school dropout status. If a person has not completed high school or obtained a GED certificate, we assign $k_1 = 1$. Otherwise we assign $k_1 = 2, \ldots, 5$ based on the 12th grade test score reported in the 2004 interview, with each quartile as a cutoff. The 12th grade test score is missing for 1,769 high school graduates, so we impute their test score using the 10th grade math and reading score, demographic characteristics, and the state of residence. The linear prediction model for the imputation is estimated using the 12,374 high school graduates whose 12th grade scores are available. The estimated linear model has an adjusted R^2 of 0.803.

We determine college choice o_2 based on the college enrollment history reported in the 2006 and 2012 interviews. The type of the first college enrollment is reported in both 2006 and 2012 interviews, while type of the last college enrollment is reported in the 2012 interview. We assign $o_2 = 0$ if a person never enrolled in any college. For the other cases, we use the first college choice to determine $o_2 \in \{1, 2, 3\}$, except that we use the last college if a student first enrolls in a 2-year college and then goes to a 4-year college. We assign $o_2 = 1$ if a person enrolled in a 2-year college, $o_2 = 2$ if a person enrolled in a 4-year in-state public college, and $o_2 = 3$ if a person enrolled in a 4-year out-of-state public college or 4-year private college. Table B2 summarizes the number of observations for each actual enrollment pattern.

Table B2: College Choices

Enrollment pattern	Observations		Assigned o_2
Never enrolled in college	1,820	13.3%	0
Enrolled in college, no further info	182	1.3%	treated as missing
2-year college	3,483	25.5%	1
4-year public in-state	3,639	26.7%	2
2-year to 4-year public	883	6.5%	2
4-year public out-of-state	666	4.9%	3
4-year private	2,454	18.0%	3
2-year to 4-year private	516	3.8%	3
Total	13,643		

We use the degree completion status reported in the 2012 interview to determine college outcome k_2. For a 2-year college student, we assign $k_2 = 2$ if the student gets any degree or certificate and $k_2 = 1$ otherwise. For a 4-year college student, we assign $k_2 = 3$ if the student graduates with a Bachelor's degree and $k_2 = 1$ otherwise. We treat the college outcome as missing for a 2-year college student whose first enrollment is in 2010 or later and for a 4-year college student whose first enrollment is in 2007 or later, since it is likely that their final college history is incomplete. We use the outstanding college loan level in the 2006 interview to measure the annual borrowing level d during college. To compute the annual borrowing, we divide the outstanding debt by 2 for those who enrolled in college in 2004 or earlier. We use the outstanding debt as the annual borrowing for those who first enrolled in 2005.

To estimate the aid function, we use both self-reported aid information as well as the Pell grant received in the first year of college. For each college student, we first impute the aid amount $\tilde{A}$ as

$$\tilde{A} = \max\left\{(\text{Pell Grant}), \alpha \times (\text{Tuition})\right\},$$

where the fraction α is determined based on self-reported information. We assign $\alpha = 1$ if

the report indicates that all college costs are covered by the aid, $\alpha = 2/3$ if at least half but not all of the cost is covered, $\alpha = 1/3$ if less than half of the cost is covered, and $\alpha = 0$ if none of the cost is covered. We then perform a Tobit regression of $\tilde{A}$ on (x, k_1) and tuition to estimate the aid function $A(\cdot)$, as described in detail in Section 2.1 below.

1.1.2 ACS 2002

We create two subsamples from ACS 2002. The first contains all primary school students, which we use to get information on the private primary school attendance rate. The second contains pairs of siblings, where a younger sibling goes to primary school and an older sibling goes to high school. We use the to get information on the joint distribution of private attendance at primary school level and at high school level. In these subsamples, we define household type x in the same way as in the ELS 2002 sample.

1.1.3 Census of Population 2000

We use the sample of women aged 35-40 and single men aged 37-42 to estimate the state-specific demographic distribution $F_s(x)$, the fertility rate $q_s(x)$, the federal income tax schedule $\tau^0(x_1)$, and the standard local tax schedule $\tau^b(x_1)$. As we have 20 discrete household types in our model, we estimate the demographic distribution $F_s(x)$ nonparametrically. We use the fraction of households with any child as the fertility rate $q_s(x)$ for a state-type cell with 400 or more observations. We run a regression of child dummy on state dummies and household type dummies in an additively separable specification to estimate the fertility rate for the other cells with fewer observations.

To estimate the tax schedules, we impute the federal and state income tax liabilities of each household using the Internet TAXSIM (v27) program provided by NBER. We use 2002 as the tax year, converting reported income from 2000 dollars into 2002 dollars. As the federal tax schedule $\tau^0(x_1)$, we use the median tax rate of each income group. For households in the lowest income group, we also take SNAP benefits into account. While

the median federal income tax rate is 0.00% in this group, they receive $639 of SNAP benefits per household in ACS2005, which is 5.82% of their median income. So we use -5.82% as the federal income tax rate for the lowest income group.

We consider both income tax and other taxes as the sources of local tax. In addition to the imputed state income tax rate τ_{is}^I of household i in state s, we also obtain the tax rate τ_s^N associated with other local taxes, assuming a proportional tax schedule. We define τ_s^N to be the ratio of total local tax revenue excluding personal income tax (from Census of Governments 2002) to the total personal income of the state (from the U.S. Bureau of Economic Analysis). Using the total local tax rate $\tau_{is} = \tau_{is}^I + \tau_s^N$, we predict a standard tax schedule by estimating a regression

$$\log \tau_{is} = \theta_s + d_i'\gamma + \epsilon_{is},$$

where θ_s is the state fixed effect, and d_i is the vector of income category dummies. The total tax rate τ_{is} is top-coded at 50% and bottom coded at 1% in the regression. We use $\exp(d_i'\hat{\gamma})$ as the baseline local tax schedule. Table B3 presents the estimated federal and local tax schedules.

Table B3: Tax Schedules

Income group	Federal tax schedule $\tau_1^0(x_1)$	Local tax baseline $\tau^b(x_1)$
$0-20,000	-5.82%	1.0 (normalized)
$20,001-35,000	3.95%	1.150
$35,001-50,000	6.85%	1.209
$50,001-75,000	9.02%	1.261
$75,000-	15.17%	1.350

1.1.4 Other Data Sources

We use IPEDS 2002 to compute college attendance costs. For the 2-year cost, we take a weighted average of in-state tuition and fees among 2-year public colleges in each state,

where weights are in-state freshmen cohort size. For the public 4-year cost, we take a weighted average of in-state tuition and fees among 4-year public colleges in each state, where weights are in-state freshmen cohort size. For the private 4-year cost, we take a weighted average of out-of-state tuition and fees among 4-year public colleges and in-state tuition and fees among 4-year private colleges in the entire United states, where weights are out-of-state freshmen cohort size for public colleges and freshmen cohort size for private colleges.

We use the regional enrollment-weighted average private school tuition in SASS 2010-2011 for K-12 cost of attendance. Due to a rapid growth of tuition levels, we convert 2010 tuition levels into 2002 levels using the following formula, instead of merely adjusting for inflation.

$$t_{2002,r} = t_{2010,r} \cdot \frac{\bar{t}_{2003}}{\bar{t}_{2010}} \cdot \left(\frac{\bar{t}_{1999}}{\bar{t}_{2003}} \right)^{1/4}$$

$t_{y,r}$ is a regional average in year y and $\bar{t}_y$ is a national average in year y. This method gives \$7,683 for NorthEast, \$4,062 for MidWest, \$6,331 for South, and \$6,844 for West.

1.2 Government Data

1.2.1 Policy Grid

Table B4: Policy Grid

Tax rate (%)	8.5	10.0	10.3	10.8	11.5	12.5	17.0	
K-12 exp (\$1,000/yr)	5.0	6.4	6.7	7.2	7.9	8.5	10.0	12.5
College exp (\$1,000/yr)	10.5	12.0	13.5	14.5	15.3	16.1	17.2	20.5
2-year tuition (\$1,000/yr)	0.0	1.4	1.9	2.3	2.7	3.7	4.7	
4-year tuition (\$1,000/yr)	0.0	2.0	2.8	3.4	4.0	4.8	6.5	8.0

1.2.2 Census of Governments 2002

Using Census of Governments 2002, we compute the data analogues of tax revenue $\int \tau_1(x_1)\, dF(x)$, endowment z_1, K-12 expenditure e_1, net college expenditure $N_{2l}(\varphi e_2 - \underline{t}) + N_{2h}(e_2 - t)$, and expenditure on other public goods g. To obtain per-household expenditure and revenue, we use the number of families in 2002 ACS as the denominator. Each type of revenue or expenditure is constructed as below, using the categories in the data.

1. Tax Revenue: sum of all tax revenue, including both income tax and other taxes
2. Outside Revenue: sum of all following non-tax revenue items
   - Net liquor store revenue
   - Net revenue from parking facilities
   - Net interest earnings
   - Net Intergovernmental revenue from federal government
   - Misc general revenue: special assessments
   - Misc general revenue: sale of property
   - Misc general revenue: other general revenue
3. K-12 Expenditure:

   [Expenditure: Elementary and Secondary] - [Current Charges: School lunch sales]
4. Net College Expenditure:

   [Expenditure: Higher Education] - [Current Charges: Institutions of higher education]
5. Expenditure on Other Public Goods:

   [Tax Revenue]+[Endowment] - [K12 Expenditure] - [Net College Expenditure]

Although state and local governments also receive some non-tax revenue from public services such as hospitals, highways, and airports, we exclude these revenues, since spending on those services typically exceeds revenue.

2 Estimation Details

2.1 Financial Aid Function

We estimate the financial aid function outside of the model. Using the imputed aid amount $\tilde{A}$ from the ELS 2002 data, we estimate a Tobit model

$$\tilde{A} = \max\left\{0, \mu_{o_2}^A\left(C, x, k_1\right) + \epsilon\right\}, \epsilon \sim N\left(0, \sigma_{o_2}^A\right),$$

for each college type $o_2 = 1, 2, 3$. We use the conditional mean of the Tobit model as the aid function $A_{o_2}(\cdot)$:

$$A_{o_2}(C, x, k_1) = \mu_{o_2}^A\left(C, x, k_1\right)\Phi\left(\mu_{o_2}^A\left(C, x, k_1\right)\right) + \sigma_{o_2}\phi\left(\mu_{o_2}^A\left(C, x, k_1\right)\right).$$

Table B5 presents the parameters of the estimated Tobit model.

Table B5: Aid Function

Parameter	2-year	4-year public	4-year private
Intercept	-0.62	1.37	5.10
Inc=2	0.14	0.01	-1.77
Inc=3	-0.27	-1.11	-1.38
Inc=4	-0.34	-1.10	-2.91
Inc=5	-0.75	-1.38	-4.52
College	-0.02	-0.27	0.31
Minority	0.31	0.63	1.33
$k_1 = 2$	-0.03	0.13	0.62
$k_1 = 3$	-0.01	0.31	0.92
$k_1 = 4$	0.06	0.86	2.94
Tuition*(Inc=1)	1.11	0.17	
Tuition*(Inc=2)	0.92	-0.01	
Tuition*(Inc=3)	0.81	0.06	
Tuition*(Inc=4)	0.47	-0.01	
Tuition*(Inc=5)	0.62	-0.08	
Tuition^2	-0.11	0.02	
σ_{o_2}	2.31	2.10	6.19

Tuition and aid are in $1,000 units.