Journal of Public Economics 49 (1992) 135–162. North-Holland

The demand for education with 'power equalizing' aid

Estimation and simulation

Paul Rothstein*

Department of Economics, Washington University in St. Louis, St. Louis, MO 63130, USA

Received May 1990, final version received August 1991

State grants-in-aid modify the full fiscal income and tax price for education services faced by voters in local school districts. If aid is non-linear in the local property tax rate, as occurs with District Power Equalization coupled with an aid floor, then individual budget constraints are non-linear and least squares estimation is inappropriate. We develop and estimate a model that incorporates these effects, derive and extensively illustrate the comparative statics, present global and local elasticities, and simulate the impact of parameter changes on the variance of spending and the cost of the program.

1. Introduction

State contributions are the largest source of revenues available to public elementary and secondary schools, followed closely by property tax revenue.[1] Given the magnitude of aid, it seems we should account for it carefully when modeling the spending choices of local school districts. This is made difficult by the fact that the rules governing the state's contribution generally make the aid positively or negatively contingent on any number of district characteristics, including the local property tax rate. If aid is linear in the tax rate, then one can use the standard models that incorporate matching and block grants [Craig and Inman (1986)] and estimate them with ordinary least squares [Borcherding and Deacon (1972), Bergstrom and Goodman

Correspondence to: P. Rothstein, Department of Economics, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130, USA.

*Part of this research appears in the author's doctoral dissertation and in a working paper presented at the 1989 meetings of the Regional Science Association in Santa Barbara, California. It was completed while the author was a post-doctoral fellow at the Graduate School of Industrial Administration, Carnegie Mellon University. I thank Art Denzau, Thomas Downes, Dennis Epple, Steve Garber, John Londregan, John Quigley, Howard Rosenthal, Dan Rubinfeld, Perry Shapiro, Dan Slesnick, and Lowell Taylor for their comments at various stages. All remaining errors are mine.

[1] Tax Foundation (1990). In 1970, state governments contributed 40 percent of revenues and communities 52 percent, while in 1987 the respective shares were 50 percent and 44 percent. Aggregate figures mask the tremendous diversity across states in these figures.

(1973), Inman (1979), Oates (1986)]. This holds, for example with the simplest District Power Equalization (DPE) formulas [Coones, Clune and Sugarman (1970), Stern (1973), Feldstein (1975), Aronson and Hilley (1976) Inman and Rubinfeld (1979), Phelps and Addonizio (1981, 1983)]. However, the simplest formulas are never used, and aid is generally non-linear in the local tax rate. This produces a number of conceptual and econometric problems that require a careful analysis.

We begin with the conceptual problems and derive the individual budget constraint when state aid is non-linear in the property tax rate. This result is then used to analyze expenditures in our sample of Michigan school districts in 1981–1982, when a modified power equalization formula was in effect. For property poor districts, the state's contribution increases linearly with the tax rate (without bound), while for property rich districts it decreases linearly to a non-negative value and then remains constant. This produces a single non-convexity for the subsample of property rich districts at the point where aid stops falling. A series of papers by Burtless and Hausman (1978), Hausman and Wise (1980), Moffitt (1984, 1986, 1989), and Hausman (1985) develop a general maximum likelihood approach for obtaining consistent estimates in this context, called the '2-error model'. One form of this model is well suited to our problem [Hausman and Wise (1980)] and we apply it to obtain consistent estimates.

We then derive, interpret, and extensively illustrate the comparative statics for the stochastic model. These consider how *expected* spending varies with changes in the parameters of the model, including parameters of the aid formula. This is done to a very limited degree in the literature on the 2-error model, no doubt because the analytical expressions become very complicated the more complicated the budget constraint. Our model is simple enough that the formulas are tractable and have a strong economic intuition. Furthermore, they provide a complete analysis of the impact of the power equalization program, something a purely numerical approach cannot accomplish.

Following the introduction we give a general expression for the individual budget constraint in the presence of state aid, as well as definitions of the marginal tax price and full fiscal income. We use these results to derive the budget constraint for any individual in a property rich district. The constraint is piecewise linear and the budget set is non-convex. Section 3 contains a model and stochastic specification that allows us to consistently estimate the parameters of this model. Technically, this adapts the Hausman and Wise (1980) model to the case of a non-convex budget set. Section 4 presents and evaluates the econometric results. Section 5 contains the comparative statics for the stochastic model and illustrates them in a number of ways with our estimates. We also consider different definitions of the income and price elasticities and illustrate the power of the model for simulating policy changes and developing recommendations. Section 6 summarizes the results.

2. Preliminaries

2.1. Tax price and full fiscal income with state aid

In each district, we group school revenues into three classes: local property taxes, 'formula aid' that may depend on the local property tax levied, and other aid. Letting Q denote district operating expenditures per pupil, we equate this with district revenues to obtain:

$$Q=mV+S(m)+F, \tag{1}$$

where m denotes the local property tax rate and V, S, and F measure the per-pupil tax base, formula aid, and other aid, respectively.

Wherever S is differentiable the relationship above implicitly defines the differentiable function $m(Q)$, which is the property tax rate that equates expenditures with revenues. Substituting $m(Q)$ back into (1) gives an identity which we differentiate to obtain $\partial m/\partial Q=(V+\partial S/\partial m)^{-1}$. Since the marginal price of Q is the extra money the individual pays to increase Q one dollar, we have

$$P^{\mathrm{M},i}\equiv\frac{\partial[m(Q)H^i]}{\partial Q}=\frac{H^i}{V+\partial S/\partial m}, \tag{2}$$

where H^i is the assessed valuation of individual i's house.

Individual i in this district faces the constraint

$$Y^i=I^i-m(Q)H^i, \tag{3}$$

where I^i is before-tax income and Y^i is expenditures on all other goods. Adding $P^{\mathrm{M},i}Q$ to both sides of (3), then using (1) and (2) to eliminate Q and H, respectively, on the right-hand side yields

$$Y^i+P^{\mathrm{M},i}Q=I^i+P^{\mathrm{M},i}\{F+S[m(Q)]-m(Q)(\partial S/\partial m)\}. \tag{4}$$

The left-hand side is that of a standard budget constraint, with the correct marginal price as the coefficient of Q. The right-hand side depends in general on the level of Q, through m. However, if state aid in linear in m, so $S(m)=d+am$, then $P^{\mathrm{M},i}=H/(V+a)$ and (4) becomes

$$Y^i + \frac{H^i}{V+a} Q = I^i + \frac{H^i}{V+a}(d+F). \tag{5}$$

The quantity on the right is now called individual i's *full fiscal income* [Craig and Inman (1986)]. This captures the impact of the lump-sum portion of aid and yields a constraint in the marginal price that is fully analogous to the standard individual budget constraint.

2.2. Power equalization

State aid per pupil is given by the basic Michigan district power equalization formula for 1981–1982:[2]

$$S(m) = \max[R + b + m(V^* - V), R/3], \tag{6}$$

where $b = \$360$ for all districts and is called the 'front end allowance', $V^* = \$50{,}550$ for all districts and is called the 'guaranteed base', and $R \geqq 0$ may vary across districts but is constant in m. $S(m)$ is never negative, so there is no 'recapture' in the formula. Combining (1) and (6) yields

$$Q = \max(c_1 + mV^*, c_2 + mV), \tag{7}$$

where

$$c_1 \equiv R + b + F > c_2 \equiv R/3 + F. \tag{8}$$

For districts with $V \leqq V^*$, state aid and district spending are always given by the first components of (6) and (7) and state aid is increasing in m. Fig. 1 illustrates (6) and (7) for districts with $V > V^*$. Expenditures at any level of property tax are given by the higher curve, and state aid is a decreasing function of the tax rate up until the cut-off point. State aid is therefore piecewise linear in m.[3]

This suggests two distinct characterizations of the financial status of school districts and both prove to be important. *High-tax-base* districts are those for which $V > V^*$ while *low-tax-base* districts have $V \leqq V^*$. *In-formula* districts are those for which $S > R/3$ while *out-of-formula* districts have $S = R/3$. It is easy to show that (i) the formula guarantees each district a property tax base of at least V^*, (ii) low-tax-base districts are always in-formula, and (iii) high-

[2]This formula is derived from a document prepared by the Michigan Department of Education (1986) and the discussion of the Michigan DPE program in papers by Phelps and Addonizio (1981, 1983). See section A.1 in the appendix.

[3]More generally, for districts with $V > V^*$, a grant on the front is necessary for the non-linearity if no recapture is allowed, since $d > 0$ implies $S(m) = \max[m(V^* - V), d] \equiv d$. The grant is no longer necessary if $d < 0$ is allowed.

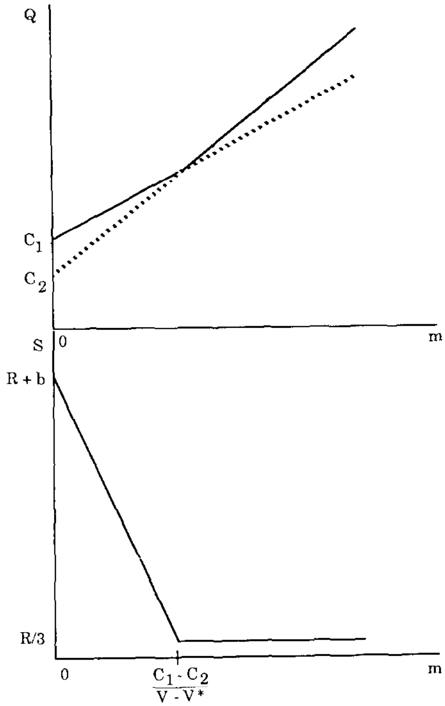

Fig. 1. Spending and aid in high-tax-base districts.

tax-base districts are in-formula if and only if spending is low enough, specifically [using (6) and (8)] $m<(c_1-c_2)/(V-V^*)$, or equivalently [using (1)] $Q<(P_1c_1-P_2c_2)/P_1-P_2)$.

The budget constraints now follow from (5). For in-formula districts $d=R+b$ and $a=(V^*-V)$ and for out-of-formula districts $d=R/3$ and $a=0$, so:

Condition	Defined	Constraint	
Low-tax-base, in-formula	$V \leqq V^*$	$Y^i+P_1Q=I^i+P_1c_1$	
High-tax-base in-formula	$V>V^*$ and $Q<\dfrac{P_1c_1-P_2c_2}{P_1-P_2}$	$Y^i+P_1Q=I^i+P_1c_1$	(9)
High-tax-base, out-of-formula	$V>V^*$ and $Q\geqq\dfrac{P_1c_1-P_2c_2}{P_1-P_2}$	$Y^i+P_2Q=I^i+P_2c_2$	

where $P_1=H/V^*$ and $P_2=H/V$.

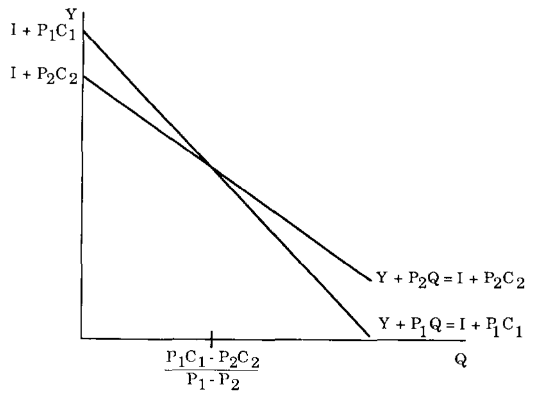

Fig. 2. Budget constraints, high-tax-base districts ($V > V^*$).

For the high-tax-base constraint, $P_1 > P_2$, so the two lines necessarily intersect. The value of Q at the intersection is the same as the spending level at which the regimes change, so the constraint is continuous. Since $c_1 > c_2$, the value of Q at the intersection is positive. Define

$$\frac{I+P_2c_2}{P_2} - \frac{I+P_1c_1}{P_1} \equiv \rho. \tag{10}$$

This is the difference in the horizontal intercepts of the two lines. The value of Y at the intersection is $\rho[(1/P_2)-(1/P_1)]^{-1}$, and this is positive if and only if ρ is positive. Therefore, the two lines intersect in the positive quadrant if and only if ρ is positive. This quantity appears repeatedly in what follows.

Fig. 2 draws the high-tax-base constraint for $\rho > 0$. It is the higher line at each value of Q. The budget set is non-convex.

3. A model of education spending

3.1. Theoretical model for high-tax-base districts

Section 2 demonstrated that the aid formula creates two regimes with different tax prices and full fiscal incomes. The high-tax-base districts ($V > V^*$) may fall into either regime depending on the level of spending they choose. Tax price and full fiscal income are therefore endogenous, conditional on the quantity of educational spending chosen.

It is well understood from the work of Hausman (1985) and Moffitt (1986) that using observed tax price and full fiscal income in the standard log-linear model and applying ordinary least squares produces biased and inconsistent estimates. The problem is that positive errors will be associated with the out-of-formula tax price and income, while negative errors will have the reverse association. The error term is therefore correlated with these independent variables. Similarly, we create selection bias by using a sample of only in-formula districts, since this omits observations with a high realization of the error term. More sophisticated estimation techniques are therefore required.

The approach here follows Hausman and Wise (1980) and extends this model by developing the comparative statics along lines described by Moffitt (1984). We suppose there is a representative individual, with preferences over education expenditures per pupil and other expenditures, who chooses the expenditures that give higher utility.[4] Denoting indirect utility in regime j by U_j^*, we then have

$$U_j^* = \max U(Q, Y)$$

$$\text{s.t. } Y + P_j Q = I + P_j c_j \tag{11}$$

for $j = 1, 2$. Preferences are given by

$$U(Q, Y) = Q^{\beta(I)} Y^{1-\beta(I)}, \quad 0 < \beta(I) < 1. \tag{12}$$

The exponent will be modelled as a function of a number of factors, but we indicate at the outset that it depends on before-tax income, I. Demand in regime j is then

$$Q_j^* = \frac{\beta(I)(I + P_j c_j)}{P_j}, \tag{13}$$

and indirect utility is

$$U_j^* = \left[\frac{\beta(I)(I + P_j c_j)}{P_j}\right]^{\beta(I)} [(1 - \beta(I))(I + P_j c_j)]^{1-\beta(I)}. \tag{14}$$

To solve for the unconditional demand, Q^*, we first equate U_1^* and U_2^* to find the value of β at which the individual is just indifferent between the regimes:

[4]Moffitt (1984) and Romer, Rosenthal and Munley (1987) use the representative individual approach in 2-error models. The former argues the need for this 'structural' approach and the latter offer interpretations in light of the agenda-setter model in Romer and Rosenthal (1979).

$$\bar{\beta}=\frac{\ln\left[\dfrac{I+P_1c_1}{I+P_2c_2}\right]}{\ln\left[\dfrac{P_1}{P_2}\right]}. \tag{15}$$

Since $c_1>c_2$ and $P_1>P_2$ for high-tax-base districts we know that $\bar{\beta}>0$. It is not constrained to be smaller than 1. However, recalling (10), $\bar{\beta}\geqq 1\leftrightarrow\rho\leqq 0$, i.e. the curves do not cross in the positive quadrant. In this case the in-formula constraint lies everywhere above the out-of-formula constraint and necessarily $Q^*=Q_1^*$.

For $\bar{\beta}<1$ (so $\rho>0$), we substitute $\bar{\beta}$ into (12) and (14) and set these equations equal to each other to obtain the indifference curve that is tangent to both legs of the budget constraint. Any lower level of β implies less of a taste for education spending and a choice of the in-formula regime, while any greater level of β implies a stronger taste for education spending and a choice of the out-of-formula regime, i.e. $\beta(I)<\bar{\beta}\leftrightarrow U_1^*>U_2^*$. Unconditional spending may then be written

$$Q^*=\begin{cases} Q_1^*, & \text{if } Q_1^*\leqq\dfrac{\bar{\beta}(I+P_1c_1)}{P_1}, \\ Q_2^*, & \text{if } Q_2^*\geqq\dfrac{\bar{\beta}(I+P_2c_2)}{P_2}. \end{cases} \tag{16}$$

Recalling the definition of Q_j^* in (13), we know Q^* is single-valued (except when $\beta=\bar{\beta}$, which is immaterial in the stochastic model), $Q_1^*<Q_2^*$ by $\rho>0$, and Q^* cannot take values between the bounds defined in (16).[5]

3.2. Stochastic specification

The stochastic specification is an application of the canonical 2-error model. We first suppose that β is a linear function of a vector of

[5]This formulation of preferences follows that in Hausman and Wise (1980). It is reasonable to ask whether the CES formulation in Burtless and Hausman (1978) is not obviously superior. Since one goal here is to develop analytical results, it is essential to have the closed-form expression for $\bar{\beta}$ in (15). This does not exist for the CES model [see eq. (14) in Burtless and Hausman]. Also, we provide a version of the model at the end of the next section with flexibility in the price elasticity [this follows eq. (4.1) in Hausman and Wise], while eq. (12) already gives flexibility to the income elasticity. The model of expenditures that results is therefore different from the CES model but not overly restrictive. Furthermore, the stochastic structures of the two models are the same [compare (22) with eq. (16) in Burtless and Hausman], except that we impose no restrictions on the coefficient.

characteristics, X, and a normally distributed random variable, η, reflecting unobserved attributes. This is the 'heterogeneity error':

$$\beta(I)=X'\delta'+I\delta_I+\eta=X\delta+\eta, \quad \eta\sim N(0,\sigma_\eta^2). \tag{17}$$

No restrictions are imposed a priori. The density for β is then $f(\beta)=N(X\delta,\sigma_\eta^2)$ and the density (up to a set of measure zero) for Q^* is

$$f(Q^*)=\begin{cases} \dfrac{1}{\sigma_1}\phi\left(\dfrac{Q^*-\mu_1}{\sigma_1}\right), & \text{if } Q^*<\dfrac{\bar{\beta}(I+P_1c_1)}{P_1}, \\ \dfrac{1}{\sigma_2}\phi\left(\dfrac{Q^*-\mu_2}{\sigma_2}\right), & \text{if } Q^*>\dfrac{\bar{\beta}(I+P_2c_2)}{P_2}, \\ 0, & \text{otherwise,} \end{cases} \tag{18}$$

where

$$\begin{aligned} \mu_1&=\frac{X\delta(I+P_1c_1)}{P_1}, & \sigma_1&=\frac{\sigma_\eta(I+P_1c_1)}{P_1}, \\ \mu_2&=\frac{X\delta(I+P_2c_2)}{P_2}, & \sigma_2&=\frac{\sigma_\eta(I+P_2c_2)}{P_2}. \end{aligned} \tag{19}$$

For each observation there is an interval where the density is indentically zero for any parameter vector δ,σ_η. If observed spending falls in this zone, the likelihood of this observation, and therefore of the entire sample, is identically zero [see also Moffitt (1986)]. This motivates introducing the second error:

$$Q=Q^*+\varepsilon, \quad \varepsilon\sim N(0,\sigma_\varepsilon^2). \tag{20}$$

Now, for any realization of η and any spending level in the zero-interval, there is a corresponding value of ε which makes the likelihood of the observation non-zero.

To find the density for Q, we find the joint density of Q and β and then integrate over β. On the half-plane given by $\beta<\bar{\beta}$, the two random variables can be written as the following affine transformations of two independent standard normal random variables:

Table 1

Parameters and densities.

Parameter	Distribution	Definition
$\mu_1, \sigma_1^2, \rho_1$	$f_1(Q) = N(\mu_1, \sigma_1^2)$	$\mu_1 = \dfrac{X\delta(I + P_1 c_1)}{P_1}$
	$f_1(Q, \beta) =$	$\sigma_1^2 = \sigma_\eta^2 \left(\dfrac{I + P_1 c_1}{P_1}\right)^2 + \sigma_\varepsilon^2$
	$BN(\mu_1, X\delta, \sigma_1^2, \sigma_\eta^2, \rho_1)$	$\rho_1 = \dfrac{I + P_1 c_1}{P_1} \left(\dfrac{\sigma_\eta^2}{\sigma_1^2}\right)^{1/2}$
θ_1, ω_1^2	$f_1(\beta \mid Q) = N(\theta_1, \omega_1^2)$	$\theta_1 = X\delta + \rho_1(\sigma_\eta/\sigma_1)(Q - \mu_1)$ $\omega_1^2 = \sigma_\eta^2(1 - \rho_1^2)$
$\mu_2, \sigma_2^2, \rho_2$	$f_2(Q) = N(\mu_2, \sigma_2^2)$	$\mu_2 = \dfrac{X\delta(I + P_2 c_2)}{P_2}$
	$f_2(Q, \beta) =$	$\sigma_2^2 = \sigma_\eta^2 \left(\dfrac{I + P_2 c_2}{P_2}\right)^2 + \sigma_\varepsilon^2$
	$BN(\mu_2, X\delta, \sigma_2^2, \sigma_\eta^2, \rho_2)$	$\rho_2 = \dfrac{I + P_2 c_2}{P_2} \left(\dfrac{\sigma_\eta^2}{\sigma_2^2}\right)^{1/2}$
θ_2, ω_2^2	$f_2(\beta \mid Q) = N(\theta_2, \omega_2^2)$	$\theta_2 = X\delta + \rho_2(\sigma_\eta/\sigma_2)(Q - \mu_2)$ $\omega_2^2 = \sigma_\eta^2(1 - \rho_2^2)$

$$Q = \frac{X\delta(I + P_1 c_1)}{P_1} + \left(\frac{I + P_1 c_1}{P_1}\right)\sigma_\eta \left(\frac{\eta}{\sigma_\eta}\right) + \sigma_\varepsilon \left(\frac{\varepsilon}{\sigma_\varepsilon}\right),$$
$$\beta(I) = X\delta + \sigma_\eta \left(\frac{\eta}{\sigma_\eta}\right). \tag{21}$$

It is well known [Bickel and Doksum (1977, Theorem 1.4.1)] that Q and β are therefore bivariate normal with means, variances, and correlation determined by the constants in (21); see table 1. Call this density $f_1(Q, \beta)$. On the half-plane $\beta > \bar{\beta}$ we have the analogous density, $f_2(Q, \beta)$. The joint density of Q and β is then

$$f(Q, \beta) = \begin{cases} f_1(Q, \beta), & \text{if } \beta \leq \bar{\beta} \\ f_2(Q, \beta), & \text{if } \beta > \bar{\beta} \end{cases}$$

(we make an arbitrary assignment on the line $\beta = \bar{\beta}$, which has measure zero in R^2). This is obviously a valid density since integrating over Q returns the density for β. Integrating over β yields the density for Q:

$$l(Q)=\int_{-\infty}^{\bar{\beta}} f_1(Q,\beta)\,\mathrm{d}\beta+\int_{\bar{\beta}}^{\infty} f_2(Q,\beta)\,\mathrm{d}\beta$$

$$=f_1(Q)\int_{-\infty}^{\bar{\beta}} f_1(\beta|Q)\,\mathrm{d}\beta+f_2(Q)\int_{\bar{\beta}}^{\infty} f_2(\beta|Q)\,\mathrm{d}\beta$$

$$=\frac{1}{\sigma_1}\phi\left(\frac{Q-\mu_1}{\sigma_1}\right)\Phi\left(\frac{\bar{\beta}-\theta_1}{\omega_1}\right)+\frac{1}{\sigma_2}\phi\left(\frac{Q-\mu_2}{\sigma_2}\right)\left[1-\Phi\left(\frac{\bar{\beta}-\theta_2}{\omega_2}\right)\right]. \quad (22)$$

To keep the notation from proliferating we redefine σ_1 and σ_2 from (19) to incorporate the term σ_ε. Throughout, ϕ indicates the standard normal density and Φ its cumulative distribution function.[6]

We can also explore a model that is slightly more general in its treatment of the price effect of state aid. Modify $\mu_1=[X\delta(I+P_1c_1)]/P_1$ into $\mu_{1,\gamma}=[X\delta(I+P_1c_1)]/P_1^\gamma$ (with similar changes to obtain $\sigma^2_{1,\gamma},\rho_{1,\gamma},\theta_{1,\gamma}$, and $\omega^2_{1,\gamma}$). The parameter γ may be interpreted as a 'discount factor' or 'perception error' in the pure price effect. For example, the fact that the guaranteed base changes most years and is therefore transitory may cause people to misperceive or discount it. In any case, this adds flexibility to the model and we can test the restriction $\gamma=1$.

3.3. *Model and stochastic specification for low-tax-base districts and the pooled sample*

For low-tax-base districts the derivation of the density of Q is straightforward. Demand is given by Q_1^* in (13) and the stochastic structure follows from (17) and (20):

$$l(Q)=\frac{1}{\sigma_1}\phi\left(\frac{Q-\mu_1}{\sigma_1}\right). \quad (23)$$

This model is linear in the parameter vector δ but the error term is

[6]The density in (22) is closely related to the sum of eqs. (1.13) and (1.15) in Hausman and Wise (1980). Since their derivation applies to a convex constraint, there are two distinct terms corresponding to $\bar{\beta}$, giving the range of values at which the individual locates at the kink, and a third term in the likelihood function giving the likelihood and probability of locating at the kink.

Table 2

	In-formula	Out-of-formula	Total
High-tax-base	66	82	148
Low-tax-base	229	0	229
Total	295	82	377

heteroscedastic.[7] We consider the more general version of the price elasticity for this model as well.

For the pooled dataset, the likelihood is given by (22) or (23) depending on whether the observation is high-tax-base or low-tax-base, respectively.

4. Economic results

Our dataset combines information from the 1980 Census summary tapes for Michigan school districts and the Michigan Department of Education 1981–1982 Annual School District Financial Reports. Districts that could not clearly be matched between the two sources were eliminated. This yielded a sample of 377 school districts, of which 148 are high-tax-base and 229 are low-tax-base (see table 2).[8]

Table 3, part (a) shows the gap in the fiscal capacity of high-tax-base and low-tax-base districts. The latter average 3,829 students per district with an average median housing value of \$35,643, while the former districts average 2,295 students with a median housing value of \$39,426. This translates into a striking difference in the tax base per pupil: \$39,011 for low-tax-base districts versus \$75,805 for high-tax-base districts. District power equalization narrows this gap by raising the tax base for the former group to \$50,550, but this still leaves the tax base 50 percent larger for the latter. The net result is that spending per pupil is on average \$250 higher in the high-tax-base districts. In contrast, the aggregate information in table 3, part (b) reveals no clear pattern of demographic differences between the two sets of districts.

Our maximum likelihood results for high-tax-base, low-tax-base, and all districts combined are in table 4, parts (a)–(c).[9] We used least squares estimates for the low-tax-base districts as starting values for the maximum likelihood estimation of (23) and (22), although for the latter these were poor and were subsequently modified. Convergence is obtained in all cases and

[7]One may estimate it with least squares using a correction detailed in Kmenta (1986, pp. 285–286).

[8]The dataset is described more extensively in section A.2 of the appendix.

[9]Six observations in our sample have $\bar{\beta}$ larger than 1, and for these the likelihood is given by (23). Since so few observations are involved we do not treat them in any special manner. Moreover, at the solution, $X\delta$ and σ_η are sufficiently smaller than 1 that $\bar{\beta}>1 \rightarrow \Phi(\cdot) \approx 1$, which implies (22) reduces to (23) at these observations. Their contribution to the likelihood at the solution is therefore identical to what their contribution would be if we had used (22).

Table 3

(a) Means of finance variables.

	In	Out	High	Low	All
Criterion	$S>R/3$	$S=R/3$	$V>50{,}550$	$V\leqq 50{,}550$	
Number of districts	295	82	148	229	377
Spending					
Q	2,015	2,421	2,252	2,007	2,103
	(265)	(443)	(418)	(265)	(354)
Number of students	3,606	1,864	2,295	3,829	3,227
	(13,016)	(2,942)	(3,097)	(14,672)	(11,612)
Median house value					
H	36,535	39,262	39,426	35,643	37,128
	(10,467)	(14,217)	(13,429)	(9,657)	(11,423)
Tax base					
V	42,667	92,267	75,805	39,011	53,455
	(9,377)	(31,335)	(29,756)	(7,060)	(26,459)
Median price					
$P_1=H/V^*$	0.72275	0.77670	0.77995	0.70511	0.73449
	(0.20707)	(0.28124)	(0.26566)	(0.19104)	(0.22597)
$P_2=H/V$	0.88487	0.46827	0.58063	0.93232	0.79426
	(0.26772)	(0.23151)	(0.26379)	(0.25822)	(0.31178)
Median family income					
I	20,408	19,228	19,943	20,286	20,152
	(4,214)	(5,954)	(5,533)	(4,012)	(4,665)
Formula aid					
S	662	23	144	767	523
	(276)	(16)	(154)	(212)	(360)
Local revenue					
mV	1,237	2,292	1,986	1,130	1,466
	(332)	(440)	(506)	(266)	(564)
Fixed aid					
F	116	106	121	110	114
	(159)	(251)	(201)	(169)	(182)
Categorical					
R	75	68	71	75	74
	(38)	(47)	(43)	(38)	(40)

Table 3 (continued)
(b) Means of characteristic variables.

	In	Out	High	Low	All
Criterion	$S>R/3$	$S=R/3$	$V>50{,}550$	$V\leqq 50{,}550$	
Number of districts	295	82	148	229	377
OWNOC	0.8088	0.8057	0.8054	0.8099	0.8081
	(0.0748)	(0.0788)	(0.0801)	(0.0722)	(0.0753)
NONWT	0.0429	0.0386	0.0368	0.0453	0.0420
	(0.0890)	(0.0710)	(0.0621)	(0.0975)	(0.0854)
OLD	0.1514	0.2066	0.1842	0.1499	0.1634
	(0.0505)	(0.0652)	(0.0636)	(0.0509)	(0.0586)
POOR	0.0592	0.0658	0.0623	0.0596	0.0607
	(0.0269)	(0.0353)	(0.0325)	(0.0266)	(0.0290)
SFAM	0.8253	0.6837	0.7222	0.8412	0.7945
	(0.1216)	(0.1153)	(0.1228)	(0.1189)	(0.1336)
URB	0.2054	0.1930	0.2264	0.1874	0.2027
	(0.3709)	(0.3804)	(0.3997)	(0.3540)	(0.3725)
PRIV	0.0604	0.0779	0.0740	0.0578	0.0642
	(0.0586)	(0.0810)	(0.0749)	(0.0559)	(0.0644)

Notes: *OWNOC*: fraction of occupied housing occupied by owner; *NONWT*: fraction of individuals non-white; *OLD*: fraction over 65; *POOR*: fraction with income below \$5,000; *SFAM*: number of students per family; *URB*: fraction of population in an urbanized areas (at least 50,000 people with 1,000 people per square mile); *PRIV*: children in private kindergarten, elementary, and high schools as fraction of children age 5–17.

small changes in the starting values, a few parameters at a time, lead to identical solutions. However, the rate of convergence is very sensitive to the starting values, and large perturbations or poor scaling of the variables can cause the procedure to fail to converge. These problems seem to be common with this model [Moffitt (1986, 1989)]. No non-converging routines showed likelihood values superior to those in the respective converging routines. Techniques discussed in Cramer (1986) proved helpful.

The implications of the estimates are best illustrated through the comparative statics and simulations in the next section. Here we briefly consider questions of fit and specification. First, although we do not constrain β in the estimation, we find that *all* of the $(4)(148)+(4)(229)+(4)(377)=3{,}016$ predicted values (one prediction for each observation in each of the 12 models) lie strictly between 0 and 1. Next, compare the predicted and actual values for β and Q. Using the estimates from the second column of table 4, part (a), we calculate the predicted value of β for each observation in the sample and then use the correct price and income term (depending on whether the district is in-formula or out-of-formula) to predict Q. The means (and

Table 4

(a) Maximum likelihood estimates for high-tax-base districts ($V > V^*$).

Constant	0.0592[a]	0.0682[a]	0.0534[a]	0.0617[a]
	(0.0202)	(0.0207)	(0.0190)	(0.0182)
OWNOC	0.0050	0.0049	−0.0135	0.0058
	(0.0161)	(0.0167)	(0.0158)	(0.0130)
NONWT	−0.0252	−0.0295	−0.0264	–
	(0.0176)	(0.0184)	(0.0187)	
OLD	−0.0884[a]	−0.1013[a]	−0.0681[a]	−0.0847[a]
	(0.0253)	(0.0259)	(0.0210)	(0.0249)
POOR	0.0443	0.0421	0.0527	–
	(0.0559)	(0.0585)	(0.0574)	
SFAM	−0.0192	−0.0179	–	−0.0219
	(0.0126)	(0.0130)		(0.0118)
URB	0.0051	0.0068	–	–
	(0.0045)	(0.0045)		
PRIV	−0.0654[a]	−0.0650[a]	–	−0.0667[a]
	(0.0167)	(0.0176)		(0.0170)
I	0.0149[a]	0.0118[a]	0.0145[a]	0.0154[a]
	(0.0045)	(0.0045)	(0.0032)	(0.0032)
γ	1.2049[a]	[1]	1.2066[a]	1.2274[a]
	(0.0792)		(0.0811)	(0.0800)
σ_η	0.0083[a]	0.0086[a]	0.0094[a]	0.0083[a]
	(0.0030)[b]	(0.0030)[b]	(0.0030)[b]	(0.0030)[b]
σ_ε	0.3090[a]	0.3085[a]	0.3022[a]	0.3183[a]
	(0.0493)	(0.0472)	(0.0475)	(0.0490)
log likelihood	−95.768	−99.133	−103.725	−97.557

$N = 148$.

[a]Significant at the 5 percent level.

[b]Value represents an upper bound. (Asymptotic standard errors.)

(b) Maximum likelihood estimates for low-tax-base districts ($V \leqq V^*$).

Constant	0.1549[a]	0.0953[a]	0.1564[a]	0.1786[a]
	(0.0145)	(0.0152)	(0.0148)	(0.0141)
OWNOC	−0.0393[a]	−0.0632[a]	−0.0568[a]	−0.0751[a]
	(0.0110)	(0.0114)	(0.0105)	(0.0095)
NONWT	0.0139	−0.0147	0.0258[a]	–
	(0.0100)	(0.0089)	(0.0100)	
OLD	0.0394	−0.0210	0.0308	0.0114
	(0.0207)	(0.0184)	(0.0205)	(0.0205)
POOR	0.1216[a]	0.0054	0.0825	–
	(0.0499)	(0.0454)	(0.0495)	
SFAM	0.0013	0.0080	–	0.0004
	(0.0063)	(0.0063)		(0.0063)
URB	0.0128[a]	0.0018	–	–
	(0.0030)[b]	(0.0030)[b]		
PRIV	−0.0069	0.0099	–	0.0111
	(0.0122)	(0.0134)		(0.0122)
I	−0.0224[a]	0.0088[a]	−0.0150[a]	−0.0164[a]
	(0.0045)	(0.0032)	(0.0045)	(0.0032)
γ	0.1389[a]	[1]	0.2390[a]	0.3312[a]
	(0.0534)		(0.0557)	(0.0522)
σ_η	0.0000	0.0000	0.0026	0.0033
	(0.0045)	(0.0032)	(0.0045)	(0.0045)
σ_ε	0.1949[a]	0.2816[a]	0.2021[a]	0.2031[a]
	(0.0089)	(0.0130)	(0.0327)	(0.0396)
log likelihood	49.512	−34.724	31.827	24.414

$N = 229$.

[a]Significant at the 5 percent level.

[b]Value represents an upper bound. (Asymptotic standard errors.)

Table 4 (continued)
(c) Maximum likelihood estimates for all districts.

Constant	0.0919[a]	0.0882[a]	0.0962[a]	0.0900[a]
	(0.0130)	(0.0126)	(0.0126)	(0.0114)
OWNOC	−0.0393[a]	−0.0385[a]	−0.0429[a]	−0.0400[a]
	(0.0100)	(0.0100)	(0.0095)	(0.0084)
NONWT	−0.0117	−0.0123	−0.0092	–
	(0.0095)	(0.0089)	(0.0089)	
OLD	−0.0621[a]	−0.0598[a]	−0.0700[a]	−0.0616[a]
	(0.0161)	(0.0158)	(0.0138)	(0.0161)
POOR	−0.0068	−0.0035	0.0127	–
	(0.0392)	(0.0385)	(0.0387)	
SFAM	0.0072	0.0054	–	0.0031
	(0.0063)	(0.0063)		(0.0063)
URB	0.0048	0.0042	–	–
	(0.0032)	(0.0030)[b]		
PRIV	−0.0190	−0.0197	–	−0.0197
	(0.0110)	(0.0110)		(0.0110)
I	0.0059	0.0074[a]	0.0077[a]	0.0086[a]
	(0.0032)	(0.0032)	(0.0032)	(0.0032)[b]
γ	0.9514[a]	[1]	0.9734[a]	0.9650[a]
	(0.0427)		(0.0420)	(0.0421)
σ_η	0.0093[a]	0.0090[a]	0.0093[a]	0.0092[a]
	(0.0030)[b]	(0.0030)[b]	(0.0030)[b]	(0.0030)[b]
σ_ε	0.2175[a]	0.2255[a]	0.2209[a]	0.2217[a]
	(0.0247)	(0.0243)	(0.0249)	(0.0251)
log likelihood	−167.306	−167.948	−171.945	−169.816

$N=377$.
[a]Significant at the 5 percent level.
[b]Value represents an upper bound. (Asymptotic standard errors.)

standard deviations) are 0.0625 (0.0107) for β and 2,292 (685) for Q. The actual value of β for each observation follows from the actual value of Q and the correct price and income terms. The means (and standard deviations) are 0.0641 (0.0166) for β and 2,252 (418) for Q, which are quite similar.

The models in columns 2–4 are nested within that in column 1. This uses all of the demographic variables and the price term P_1^γ discussed at the end of subsection 3.2. If individuals correctly perceive the guaranteed base, then $\gamma=1$. The point estimate is 1.2 and the likelihood ratio test rejects (at the 5 percent level) the restriction to 1.[10] This implies that high-tax-base districts slightly under-perceive the in-formula price (and so over-perceive the guaranteed base).[11] The model in the third column excludes three less commonly used variables, but we can reject these restrictions as well. The last model

[10]The likelihood ratio statistic equals $6.73=2(-95.768+99.133)$ while the critical value of $\chi^2(1)$ is 3.841.

[11]For high-tax-base districts we have on average $P_1=0.780$, so $(P_1)^\gamma<P_1$.

excludes three of the insignificant variables from the first model, and this restriction we cannot reject.[12]

We now briefly consider the estimates for the low-tax-base districts. Using the estimates from the second column of table 4, part (b), we find the average predicted values of β and Q are 0.0662 (0.0059) and 1,984 (294), respectively, while the average sample values are 0.0679 (0.0117) and 2,007 (265). Again the fit is quite good. However, the point estimate of γ is now 0.139, it is significantly different from both zero and 1, and the restriction $\gamma=1$ substantially affects the estimates of δ. We therefore use only the perception error model for low-tax-base districts and the column 1 estimates. This suggests that low-tax-base districts over-perceive the tax price (under-perceive the guaranteed base), which is consistent with the notion that the state aid formula is somewhat poorly understood. All other restrictions are also rejected by the likelihood ratio test, giving substantial support to the larger model.

Last, consider the consistency of the result across the two datasets. Comparing the first column of each table, we note on the one hand that the only estimate that is statistically significant in both datasets with the opposite sign is the coefficient of I. The implications for comparative statics becomes clear in the next section. On the other hand, of the eight independent variables, only the coefficients of *POOR*, *URB*, and *PRIV* have the same sign. The correlations among them are reasonably low[13] and neither dataset is very small. The results may reflect differences in the tastes of individuals who choose to live in the different districts or the need to find more direct measures of tastes than we capture with these variables.

Further comparisons are developed in the next section.

5. Comparative statics and simulations

We begin by focusing on the high-tax-base districts. Comparative statics in the 2-error model consists of measuring the change in expected expenditures with changes in an independent variable. The density with respect to which the expectations are taken is $f(Q^*)$, given in (18). Expected expenditures are then

$$\mu_{Q^*}=\int_{-\infty}^{\bar{\beta}(I+P_1c_1)/P_1}\frac{Q^*}{\sigma_1}\,\phi\left(\frac{Q^*-\mu_1}{\sigma_1}\right)\mathrm{d}Q^*+\int_{\bar{\beta}(I+P_2c_2)/P_2}^{\infty}\frac{Q^*}{\sigma_2}\,\phi\left(\frac{Q^*-\mu_2}{\sigma_2}\right)\mathrm{d}Q^*$$

[12]The likelihood ratio statistic equals $3.578=2(-95.768+97.557)$ while the critical value of $\chi^2(3)$ is 7.815.

[13]The three highest are $\mathrm{cor}(I,POOR)=-0.81$, $\mathrm{cor}(I,OLD)=-0.66$, and $\mathrm{cor}(NONWT, OWNOC)=-0.52$.

$$
= X\delta\left(\frac{I+P_1c_1}{P_1}\right)\Phi\left(\frac{\bar{\beta}-X\delta}{\sigma_\eta}\right)
$$

$$
+X\delta\left(\frac{I+P_2c_2}{P_2}\right)\left[1-\Phi\left(\frac{\bar{\beta}-X\delta}{\sigma_\eta}\right)\right]+\rho\sigma_\eta\phi\left(\frac{\bar{\beta}-X\delta}{\sigma_\eta}\right). \qquad (24)
$$

Actually, this is the expression for expected expenditures if $\rho > 0$, which holds when we set variables to the high-tax-base means. If this condition fails, then the district *must* be in-formula. However, for any observation where $\rho \leqq 0$ we have $\bar{\beta} \geqq 1$, and given our estimates[14] this implies $\Phi(\cdot) \approx 1$ and $\phi(\cdot) \approx 0$. Substituting these values in (24) shows $\mu_{Q*} \approx X\delta(I+P_1c_1/P_1) = \mu_1$, the in-formula mean. The following comparative statics are therefore generally valid.

Fig. 3(a) illustrates the initial situation, $b=\$360$ and $V^*=\$50{,}550$. Setting other variables to their means, we determine full fiscal income and price in both regimes, which then determine $\bar{\beta}$. If $\beta<\bar{\beta}$, then the district is in-formula and ideal spending must be less than 1,134, while $\beta>\bar{\beta}$ implies ideal spending must be larger than 1,672. The density is then given by the lower distribution to the left of 1,134 and the upper distribution to the right of 1,672, which we indicate by shading. The density is zero between these values. We find $\mu_1=1{,}632$, $\sigma_1=231$, $\mu_2=2{,}404$, $\sigma_2=341$, and $\mu_{Q*}=2{,}397$.

Fig. 3(b) considers the impact of setting $b=\$700$. Formally, this directly changes c_1, which then enters $\bar{\beta}$, μ_1, and σ_1. The new density is given by the lower distribution to the left of 1,994 and the upper distribution to the right of 2,900. The upper distribution is unchanged. However, the mean of the lower distribution increases, and we now have $\mu_1=1{,}653$, $\sigma_1=234$, and $\mu_{Q*}=1{,}723$.

Notice that expected expenditures fall despite the fact that the in-formula mean must increase with the outward shift in the in-formula constaint. As fig. 3(b) shows, the drop occurs because our typical district is now more likely to shift from being a relatively high-spending out-of-formula district to being in-formula, and this shift more than offsets the upward shift in μ_1.

Fig. 3(c) considers changing the guaranteed base to \$64,000. The story is the same as that just described for the front end allowance, and we have $\mu_1=2{,}057$, $\sigma_1=291$, and $\mu_{Q*}=2{,}076$.

Now consider (24) somewhat more generally. The derivatives with b and V^* are

[14]Using mean high-tax-base values for X and estimates from table 4, part (a), column 2, we have $X\delta=0.062$ and $\sigma_\eta=0.009$.

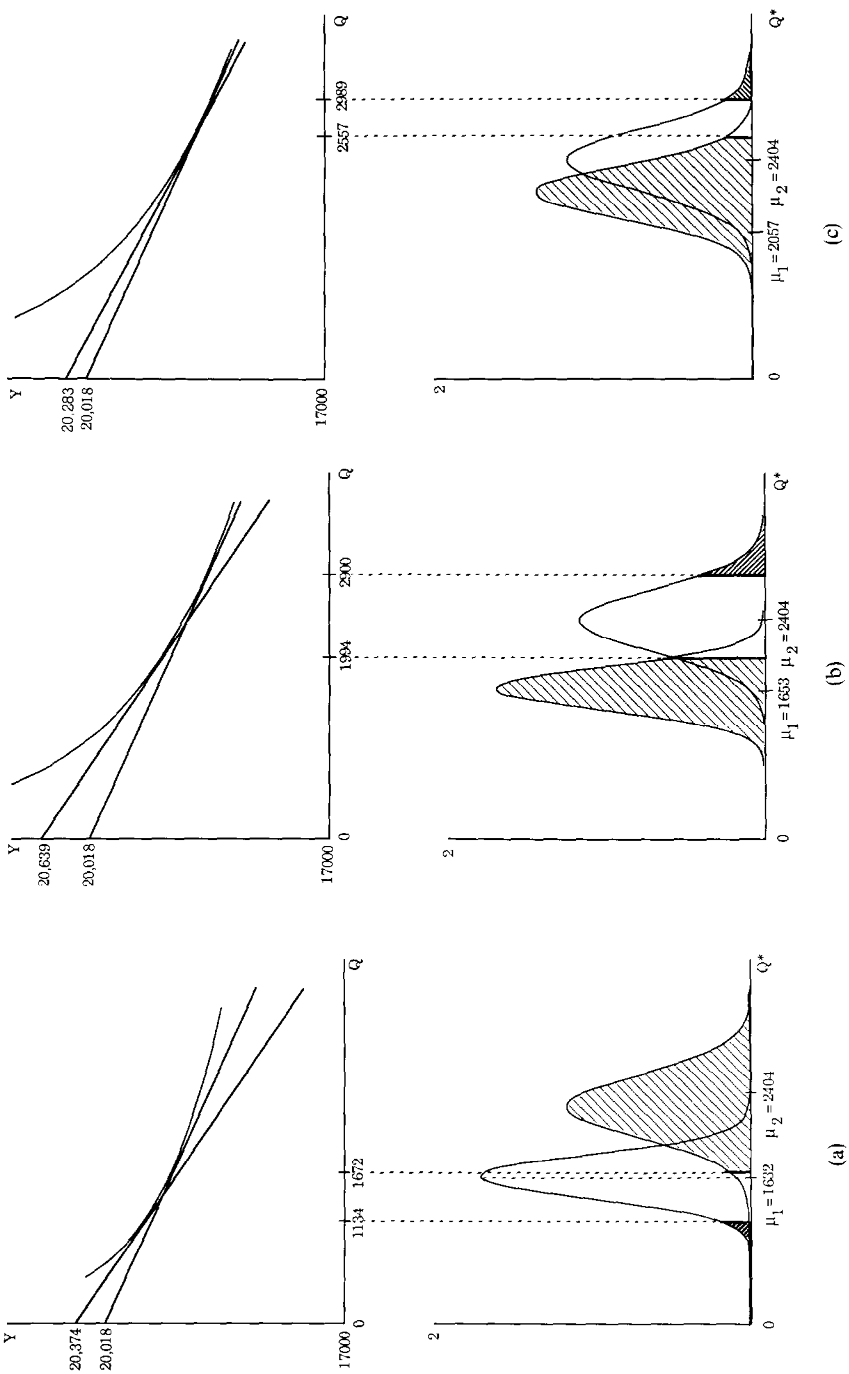

Fig. 3. (a) $b = 360$; $V^* = 50{,}550$. (b) $b = 700$; $V^* = 50{,}550$. (c) $b = 360$; $V^* = 64{,}000$.

$$\frac{\partial \mu_{Q^*}}{\partial b} = X\delta\Phi(\cdot) - \phi(\cdot)\left[\sigma_\eta + \frac{\rho P_1 \bar{\beta}}{\sigma_\eta (I + P_1 c_1)\ln(P_1/P_2)}\right], \tag{25}$$

$$\frac{\partial \mu_{Q^*}}{\partial V^*} = \frac{X\delta I\Phi(\cdot)}{H} - \frac{\phi(\cdot)}{H}\left\{I\sigma_\eta + \frac{\rho P_1 \bar{\beta}[\bar{\beta} I + (\bar{\beta}-1)P_1 c_1]}{\sigma_\eta (I + P_1 c_1)\ln(P_1/P_2)}\right\}. \tag{26}$$

Since $\rho > 0$, the term in brackets in (25) is unambiguously positive and the term in (26) is positive for $\bar{\beta}$ not too small. The derivatives are therefore negative when these terms dominate, as in the previous examples. Alternatively, for $\bar{\beta}$ relatively large, $\Phi(\cdot) \approx 1$ and $\phi(\cdot) \approx 0$, and (24) reduces to $\mu_{Q^*} = X\delta(I + P_1 c_1)/P_1 = \mu_1$. The derivatives in (25) and (26) also reduce to the respective derivatives of μ_1, and are positive. For $\bar{\beta}$ relatively small, $\Phi(\cdot) \approx 0$ and $\phi(\cdot) \approx 0$ and (24) becomes $X\delta(I + P_2 c_2)/P_2 = \mu_2$, the out-of-formula mean. The derivatives now reduce to the respective derivatives of μ_2, which are zero.

The basic intuition is that if the in-formula constraint is relatively high and flat, then the in-formula regime is relatively attractive. This occurs if b and V^* are large, and we have from (15) that $b \to \infty$ and $V^* \uparrow V$ each imply $\bar{\beta} \to \infty$. Therefore $\Phi(\cdot) \approx 1$ and the district is in-formula, as expected. If the in-formula constraint is relatively low and steep, then the in-formula regime is relatively unattractive. This occurs if b is small and V^* is moderately small, so $\bar{\beta}$ is small and $\Phi(\cdot) \approx 0$. If the in-formula constraint is steep and high, as occurs when V^* is very small, then the outcome is not as clear. However, $V^* \downarrow 0$ implies $\bar{\beta} \to 1$, so $\Phi(\cdot) \approx 1$. The increase in full fiscal income eventually dominates the increase in the slope and the district is in-formula.

These general properties are illustrated in fig. 4. In fig. 4(a), the top portion shows μ_{Q^*} as the front end allowance changes, at three values of the base. The bottom shows $\Phi(\cdot)$. $\bar{\beta}$ is monotonic and increasing in b and relatively small for small values of b (provided $V^*$is not too small). Φ therefore begins at zero and expected spending begins at the out-of-formula mean, \$2,404. As b increases, the district moves in-formula, causing spending to fall [as in fig. 3(b)]. Once the district is in-formula, spending equals the in-formula mean, and this increases slowly in b.

Fig. 4(b) shows μ_{Q^*} as the guaranteed base changes, at three values of the front end allowance. Very small values of V^* imply the district is in-formula and spending equals the in-formula mean. As V^* rises, spending increases both because the district moves out-of-formula and because the in-formula mean rises. Further increases induce the district to move back in-formula, which tends to cause spending to fall [as in fig. 3(c)]. Once the district is almost surely in-formula, $\mu_{Q^*} = \mu_1$, which increase as V^* approaches V. These

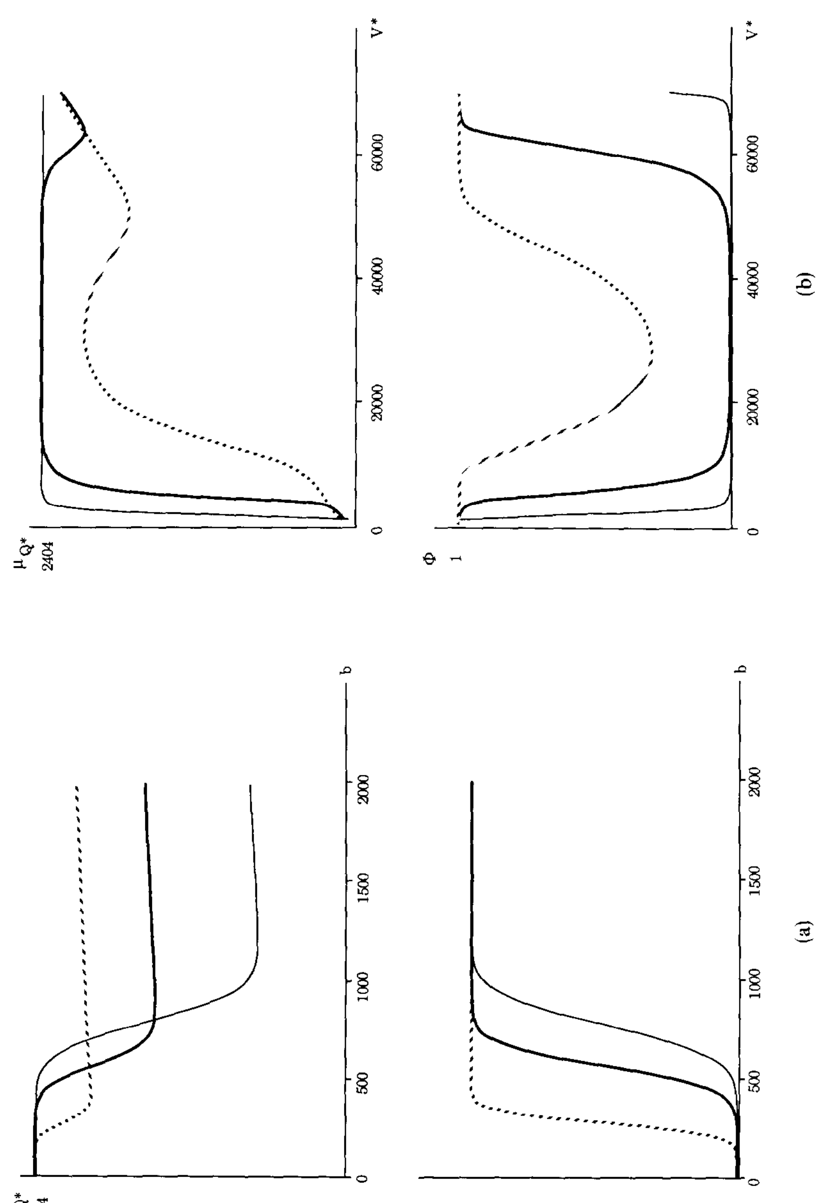

Fig. 4. (a) ---: $V^* = 64{,}000$; ——: $V^* = 50{,}550$; ——: $V^* = 30{,}000$. (b) ---: $b = 700$; ——: $b = 360$; ——: $b = 100$.

shifts are especially clear for the case $b=700$, where the in-formula probability never reaches 0 and expected spending falls as this probability rises.

We complete the comparative statics by considering income and price elasticities. These can be based on at least three measures of expenditures. The first and truly 'global' measure is μ_{Q*}. For high-tax-base districts, the derivatives are

$$\frac{\partial \mu_{Q*}}{\partial I}=\delta_I\left[\left(\frac{I+P_1c_1}{P_1}\right)\Phi(\cdot)+\left(\frac{I+P_2c_2}{P_2}\right)[1-\Phi(\cdot)]\right]$$

$$+\frac{X\delta}{P_1}\Phi(\cdot)+\frac{X\delta}{P_2}[1-\Phi(\cdot)]$$

$$-\phi(\cdot)\left[\sigma_\eta\left(\frac{1}{P_1}-\frac{1}{P_2}\right)+\frac{\rho\bar{\beta}}{\sigma_\eta}\left(\frac{1}{(I+P_1c_1)\ln(P_1/P_2)}-\delta_I\right)\right], \tag{27}$$

$$\frac{\partial \mu_{Q*}}{\partial P_1}=-\left\{\frac{X\delta I\Phi(\cdot)}{P_1^2}-\frac{\phi(\cdot)}{P_1^2}\left[I\sigma_\eta+\frac{\rho P_1\bar{\beta}(\bar{\beta}I+(\bar{\beta}-1)P_1c_1}{\sigma_\eta(I+P_1c_1)\ln(P_1/P_2)}\right]\right\}. \tag{28}$$

For low-tax-base districts, $\mu_{Q*}=\mu_1=X\delta(I+P_1c_1)/P_1$ so the derivatives are straightforward. The elasticities reported in the first two rows of table 5 use the estimates in column 2 of table 4, part (a), and the means of the independent variables.

'Local' measures of expenditures are $\mu_{1,\gamma}=X\delta(I+P_1c_1)/P_1^\gamma$ and $\mu_2=X\delta(I+P_2c_2)P_2$. For purposes of comparison we report the former elasticity for both high-tax-base and low-tax-base districts even though the typical high-tax-base district is out-of-formula (so μ_2 is more appropriate). The elasticities reported in rows 3–6 use the estimates in column 1 of table 4, part (a).

Last, elasticities (biased for high-tax-base districts) can be obtained using a log–log model and least squares:

$$\ln Q_j^*=\alpha\ln X'+\alpha_1\ln(I+P_jc_j)+\alpha_2\ln P_j+\text{error}. \tag{29}$$

However, this model is parameterized differently from the ones above. This is clear from eqs. (13) and (17) if we neglect the second (ε) error and take logs:

$$\ln Q_j^*=\ln(X'\delta'+I\delta_I+\eta)+\ln(I+P_jc_j)-\gamma\ln P_j. \tag{30}$$

The results are presented in table 5. The diversity of the estimates suggests

Table 5

Income and price elasticities.

Elasticity	High-tax-base		Low-tax-base	
$\frac{\partial \mu_{Q^*}}{\partial I} \frac{I}{\mu_{Q^*}}$	Eq. (27)	1.439	$\frac{\delta_I I}{X\delta} + \frac{I}{I+P_1c_1}$	1.252
$\frac{\partial \mu_{Q^*}}{\partial P_1} \frac{P_1}{\mu_{Q^*}}$	Eq. (28)	0.048	$\frac{P_1c_1}{I+P_1c_1}$	−0.981
$\frac{\partial \mu_{1,\gamma}}{\partial I} \frac{I}{\mu_{1,\gamma}}$	$\frac{\delta_I I}{X\delta} + \frac{I}{I+P_1c_1}$	1.466	$\frac{\delta_I I}{X\delta} + \frac{I}{I+P_1c_1}$	0.499
$\frac{\partial \mu_{1,\gamma}}{\partial P_1} \frac{P_1}{\mu_{1,\gamma}}$	$\frac{P_1c_1}{I+P_1c_1} - \gamma$	−1.184	$\frac{P_1c_1}{I+P_1c_1} - \gamma$	−0.120
$\frac{\partial \mu_2}{\partial I} \frac{I}{\mu_2}$	$\frac{\delta_I I}{X\delta} + \frac{I}{I+P_2c_2}$	0.487		
$\frac{\partial \mu_2}{\partial P_2} \frac{P_2}{\mu_2}$	$\frac{P_2c_2}{I+P_2c_2} - 1$	−0.996		
OLS[a]	α_1	0.446	α_1	0.362
	α_2	−0.322	α_2	−0.175

[a]For low-tax-base districts, OLS regression of (log) per-pupil spending on $I+P_1c_1$, P_1, *OWNOC*, *NONWT*, *OLD*, *POOR*. For high-tax-base districts, the regression uses the price and income term corresponding to whether the district is in-formula or out-of-formula.

that conceptual differences among them are important, as well as behavioral differences between the high-tax-base and low-tax-base districts. The clearest evidence of the conceptual differences comes from comparing the results for the low-tax-base districts. In the first row, the income elasticity reflects the impact on δ of the constraint $\gamma=1$, and the price elasticity reflects the Cobb–Douglas structure. In the third and fourth rows, the income elasticity changes because the estimate of δ changes substantially, and the price elasticity now depends on an estimated parameter and falls into the familiar range. The OLS coefficients are unbiased and in fact similar to those in lines 3 and 4, but the two models are structurally quite different.

For high-tax-base districts, the price elasticity in the second row is positive, reflecting the fact that in the range of the estimation, increasing P_1 causes the districts to move out-of-formula. The income elasticity is positive, in contrast to the elasticity with the front end allowance, which would be negative since increasing b causes districts to move in-formula. Evidently, the dual role of I gives it a substantially different role from b.[15] The local income elasticity is similar and the price elasticity now recovers the expected

[15]Elasticities based on the column 1 estimates are substantially the same.

negative sign. The contrast in lines 3 and 4 with the estimates for the low-tax-base districts is somewhat stiking and reflects the differences in the estimates of δ. The (biased) estimates for the high-tax-base districts are presented in the last two lines.[16]

This model can be used to derive a tremendous amount of information about policy changes and therefore to explore policy goals. We conclude by considering two interesting questions.[17] First, suppose that the state wants to minimize the variance in spending across its 377 districts by adjusting b and V^*, regardless of the cost (measured by the formula aid, S) of the resulting program. Given any proposed front end allowance and tax base, we determine the desired spending [using (14)] for each of the 377 districts and then calculate the mean and standard error. The initial values are \$2,124 with a standard error of \$570, an aggregate cost of \$728 per pupil, and 73 districts predicted out-of-formula (the actual number is 82). We then seek the minimum standard error as b varies from \$10 to \$1,000 (in increments of 10) and V^* from \$40,000 to \$100,000 (in increments of 1,000). The solution is $b=\$990$ and $V^*=\$67{,}000$, and at these values average spending is \$2,576 with standard error \$509, aggregate cost \$1,676, and 10 districts out-of-formula.

This would obviously require a dramatic reordering of state spending priorities. Suppose instead we impose the constraint that the cost of the program can increase by no more than \$100. The new solution is $b=\$30$ and $V^*=\$60{,}000$, and at these values average spending is \$2,400 with standard error \$521, aggregate cost \$817, and 83 districts out-of-formula. This suggests that raising the tax base elicits a strong response from all low-tax-base districts. In comparison with the previous result, we see that a similar level and distribution of spending is bought at a substantially lower cost.

6. Conclusion

In summary, this research develops an econometric model of school spending that incorporates the effects of a power equalization formula modified to include front end grants and exclude recapture. We derive the individual budget constraint in these circumstances, carefully deriving the tax

[16]Recent work [MaCurdy, Green and Paarsch (1988)] demonstrates that maximum likelihood estimation of the piecewise linear 2-error model with *convex* constraints implicitly constrains the parameters to obey the structure of the theoretical model. Specifically, it implicitly constrains the probability of locating at a kink point to be non-negative, and this surprisingly constrains the parameters so the Slutsky effect is correct. Their analysis does not apply to the *non-convex* constraint considered in this paper, and in fact the problem they identify does not occur with only one non-convex kink. The likelihood function (22) by construction places non-negative probability of falling in either regime. This does not mean that other problems may not be present, though.

[17]For these simulations we use the estimates for the pooled dataset in table 4, part (c), column 2.

price and full fiscal income from the community budget constraint, the individual tax constraint, and the state aid formula. This makes apparent the dependence of the structure of the individual budget constraint on the structure of the formula.

We then develop a stochastic specification and obtain maximum likelihood estimates for a number of models. A nice feature of the dataset we gather is that there is a subsample for which simple methods are appropriate and a subsample for which the fully developed 2-error model is required. The comparison illuminates some of the strengths and weaknesses of the estimates we obtain.

The comparative statics are developed to reveal the structure of this model and the impact of the aid formula. Changes in expected expenditures depend on changes in the density used in the expectation, so we begin by illustrating this for discrete changes. This explains why the derivatives change sign. Expected spending *falls* in some ranges as the tax base or front end allowance increases because the negative impact of high-spending out-of-formula districts moving in-formula overwhelms the positive impact on the in-formula mean. This lowers high-tax-base spending and narrows the variance of school district spending, which seems to be the intended qualitative effect of the non-linearity in the aid formula.

The simulations then consider the quantitative effect more precisely. They suggest that leaving some districts out-of-formula may be the most effective way to narrow the variance, and the same variance may be bought at radically different costs. Any number of policy goals can be investigated in this way.

Appendix

A.1. Derivation of the aid formula [eq. (6)]

According to Michigan Department of Education (1986), the membership aid formula in 1981–1982 is: '\$50.55 × operating mills levied, plus \$360, minus (S.E.V. per pupil × operating mills levied)'. In the notation of the paper, this is equivalent to

$$V^*m+b-Vm=b+m(V^*-V). \tag{31}$$

Eq. (31) also appears in Phelps and Addonizio (1983, p. 13). They call b the 'front end allowance' and V^* the 'guaranteed base', and we keep these definitions in the paper. They comment that by 1981 all local tax effort was reimbursable (p. 11) and that the Michigan formula has no recapture provision (p. 7). Therefore, the actual formula is not given by (31), but rather by

$$\max(b+m(V^*-V),0). \tag{32}$$

They define a district as 'in-formula' if it receives aid according to (32), and 'out-of-formula' if it does not.

This is not quite the entire story, although early versions of this work used (32) and the results are substantially the same. In 1980–1981 the legislature adopted a restriction, 'described as categorical recapture, [that] deducts from the categorical aid of out-of-formula districts an amount equal to the district's local revenue which exceeds the DPE guarantee.... The 1981–1982 act increased the recapture percentage to 66%.' As a practical matter we treat these 'categorical grants' symmetrically with the grants from (32). Denote the ex ante (i.e. before the choice of m) categorical grant by R. Then the aid formula has the following structure:

Low-tax-base districts receive aid equal to $b+m(V^*-V)>0$.

High-tax-base districts with $m<b/(V-V^*)$ receive aid equal to $b+m(V^*-V)>0$.

High-tax-base districts with $m\geqq b/(V-V^*)$ deduct the excess of local revenue over the guarantee, i.e. $mV-(b+mV^*)=-[b+m(V^*-V)]\geqq 0$, but the deduction cannot exceed $2R/3$. This is equivalent to receiving aid equal to the larger of $b+m(V^*-V)\leqq 0$ and $-2R/3$.

Since $R>0$, the entire formula is summarized by

$$\max(b+m(V^*-V),\ -2R/3). \tag{33}$$

Since we treat R symmetrically we add it to the formula, giving

$$S(m)=\max(R+b+m(V^*-V),R/3). \tag{34}$$

This is eq. (6) in the text of the paper.

We redefine 'in-formula' and 'out-of-formula' in terms of this rule. Thus, a district is in-formula if it receives more than the minimum $R/3$ (rather than 0) and out-of-formula otherwise.

A.2. Description of the dataset

The dataset combines information from the annual school financial reports prepared by the state of Michigan and the 1980 U.S. Census Summary Tape Files STF1F and STF3F for Michigan school districts. The census codes for the districts differ from the Michigan codes, so two files were created, one with Michigan codes and district names and another with Census codes and district names. The files were then merged accordingly to district names, then merged to the Michigan data by Michigan codes and to the Census data by Census codes. The initial merging was done by hand to ensure the correct districts were being matched and districts were not excluded because of different abbreviation and punctuation conventions in the datasets. The data were further merged to a Michigan dataset containing 'yes' and 'no' votes on school spending proposals, but we do not use that data in this paper.

The Michigan portion of the dataset provides Total Operating Millage (m), District State Equalized Valuation (V), State Aid Members (number of students), Total Revenue From State Sources ($S(m)$, D.P. 669), Revenue From State Sources – Restricted Grants (R, D.P. 649), and Total Current Operating Expenditures (Q, D.P. 2999). The census provides Median Family Income (I) and Median Owner Occupied Housing Value (H). From these we can determine whether a district is in-formula or out-of-formula; derive F as residual; and derive P_1, c_1, P_2, and c_2. The Census also provides all of the demographic information in the X vector.

We also verified that the theoretical relations described above are reflected fairly accurately in the dataset. In particular, Unrestricted State School Aid (D.P. 0550) closely fits (33) and Total Revenue from State Sources (D.P. 0669) closely fits (34).

References

Aronson, R. and J. Hilley, 1986, Financing state and local government, 4th edn. (The Brookings Institution, Washington, DC)

Bergstrom, T. and R. Goodman, 1973, Private demands for public goods, American Economic Review 63, 280–296.

Bickel, P. and K. Doksum, 1977, Mathematical statistics: Basic ideas and selected topics (Holden-Day, San Francisco).

Borcherding, T. and R. Deacon, 1972, The demand for the services of non-federal governments, American Economic Review 62, 891–901.

Burtless, G. and J. Hausman, 1978, The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment, Journal of Political Economy 86, 1103–1130.

Coones, J., W. Clune and S. Sugarman, 1970, Private wealth and public education (Cambridge University Press, New York).

Craig, S. and R. Inman, 1986, Education, welfare, and the 'new' federalism, in: H. Rosen, ed., Studies in state and local public finance (University of Chicago Press, Chicago, IL) 187–222.

Cramer, J.S., 1986, Econometric applications of maximum likelihood methods (Cambridge University Press, New York).

Feldstein, M., 1975, Wealth neutrality and local choice in public education, American Economic Review 65, 75–89.

Hausman, J., 1985, The econometrics of non-linear budget sets, Econometrica 53, 1255–1282.

Hausman, J. and D. Wise, 1980, Discontinuous budget constraints and estimation: The demand for housing, Review of Economic Studies 47, 75–96.

Inman, R., 1979, The fiscal performance of local governments: An interpretative review, in: P. Mieszkowski and M. Straszheim, eds., Current issues in urban economics (The Johns Hopkins Univesity Press, Baltimore, MD).

Inman, R. and D. Rubinfeld, 1979, The judicial pursuit of local fiscal equity, Harvard Law Review 92, 1662–1750.

Inter-university consortium for political and social research, 1982, Census of population and housing 1980, Michigan school districts (The Bureau of the Census, Washington, DC) summary tape files 1F and 3F.

Kmenta, J., 1986, Elements of econometrics, 2nd edition (Macmillan Publishing Company, New York).

MaCurdy, T., D. Green and H. Paarsch, 1988, Assessing empirical approaches for analyzing taxes and labor supply, typescript (Stanford University).

Michigan Department of Education, 1982, ISD and LEA: Annual school district financial report for the fiscal year ended June 30, 1982, DS-4169 Rev. Form B, on magnetic tape for all school districts.

Michigan Department of Education, 1986, The membership aid formula: Per-pupil computation, typescript.

Moffitt, R., 1984, The effects of grants-in-aid on state and local expenditures: The case of AFDC, Journal of Public Economics 23, 279–305.

Moffitt, R., 1986, The econometrics of piecewise-linear budget constraints, Journal of Business and Economic Statistics 4, 317–328.

Moffitt, R., 1989, Estimate the value of an in-kind transfer: The case of food stamps, Econometrica 57, 385–409.

Oates, W., 1986, The estimation of demand functions for local public goods: Issues in specification and interpretation, typescript (University of Maryland, MD).

Phelps, J. and M. Addonizio, 1981, District power equalizing: Cure-all or prescription, Journal of Education Finance 7, 64–87.

Phelps, J. and M. Addonizio, 1983, Michigan public school finance: The last ten years, Journal of Education Finance 9, 5–16.

Romer, T. and H. Rosenthal, 1979, Bureaucrats versus voters: On the political economy of resource allocation by direct democracy, Quarterly Journal of Economics, 563–587.

Romer, T., H. Rosenthal and V. Munley, 1987, Economic incentives and political institutions: Spending and voting in school budget referenda, Working paper 19-86-87 (Carnegie Mellon University, GSIA).

Rothstein, P., 1988, Essays in social choice theory and public sector economics, Ph.D. dissertation (University of California, Berkeley, CA).

Rothstein, P., 1989, The demand for education with power equalizing aid: Non-linear budget constraints and estimation, typescript (Carnegie Mellon University, GSIA).

Stern, D., 1973, Effects of alternative state aid formulas on the distribution of public school expenditures in Massachusetts, Review of Economics and Statistics 55, 91–97.

Tax Foundation, 1990, Facts and figures on government finance (Johns Hopkins University Press, Baltimore, MD).