An Empirical Framework for Large-Scale Policy Analysis, with an Application to School Finance Reform in Michigan

Maria Marta Ferreyra
Tepper School of Business
Carnegie Mellon University

May 30, 2008*

Abstract

In this paper I develop an empirical framework for the analysis of large-scale policies, and apply it to study the effects of school finance reform on the Detroit metropolitan area. Exploiting the school finance reform in Michigan 1994, I estimate a general equilibrium model of multiple jurisdictions with 1990 data from Detroit, predict the 2000 equilibrium, and compare this prediction with 2000 data to validate the model. I conduct counterfactual simulations using the estimates. According to my analysis, revenue-based reforms that ensure spending equity or adequacy have little impact on household demographics or school quality in Detroit. (JEL C52, I22, H73)

* Comments from the editor and two referees have substantially improved this paper. I thank the Berkman Faculty Development Fund at Carnegie Mellon University for financial support. I also thank Julie Berry Cullen and Susanna Loeb for providing me with data on pass rates for Michigan. Mary Ann Cleary, Greg Olszta and Glenda Rader from the government of the state of Michigan answered many of my questions on Proposal A and Michigan's reforms over the nineties, and facilitated my access to data. I am indebted to E. Anthon Eff, Dennis Epple, Andrew Haughwout, Holger Sieg, Fallaw Sowell and Steve Stern for valuable conversations, and to participants at the 2007 Regional Science Association meetings, the 2004 Southern Economic Association meetings, and the 2004 summer workshop at the Federal Reserve of Cleveland for comments on this and an earlier version of this paper. I thank Jeff Reminga for assistance with the technical aspects of this project, Bill Buckingham from the Applied Population Lab at University of Wisconsin-Madison for Arc GIS support, and Surendra Bagde for research assistance. All errors are mine.

1. Introduction

Although typically aiming at specific effects, large-scale policies can often trigger other, general equilibrium effects as well. Thus, a comprehensive evaluation of these policies requires a general equilibrium framework. Furthermore, it is desirable that the framework serve not only for the analysis of actual but also of counterfactual policies, particularly in light of the potentially high cost and far-reaching consequences of large-scale policies. An important issue, however, is how to develop a reliable framework. Whereas the estimation of a model provides evidence on its fit to the data, perhaps more critical evidence comes from the model's ability to fit out-of-sample data.

In this paper I develop an empirical framework for the analysis of large-scale policies that relies on a general equilibrium model, allows for counterfactual analysis and conducts model validation, and apply it to study the effects of school finance reform on the Detroit metropolitan area. Concerned about the equity and adequacy of their public school funding systems, most states have overhauled them over the last thirty years. Thus, in 1994 the state of Michigan implemented a largely unexpected reform, known as Proposal A,[1] which shifted public school funding away from local school districts onto the state. School districts no longer determine their property tax rates and hence their revenues; instead, they receive from the state a per-student allowance ("foundation allowance"). This revenue scheme increased revenue for low-revenue districts, and capped revenues for high-revenue districts. Proposal A also implemented a tax reform by reducing property tax rates on owner-occupied housing, and raising the state sales tax.

[1] See Adonizio et al (1995) and Cullen and Loeb (2004) for further details on the reform.

In metropolitan areas where households choose locations and schools jointly, a reform such as Proposal A may alter not only school revenues and property taxes, but also households' choices, housing prices, and public schools' qualities. For instance, districts that benefit from lower property taxes and higher revenues may experience an increase in property values and an improvement in local public school quality. They may also attract residents from other districts, which would further alter housing prices.

I capture these effects through an equilibrium model of multiple jurisdictions and household residential and school choice, and estimate the structural parameters using 1990 data for the Detroit metropolitan area. I use the parameter estimates to simulate the 2000 equilibrium accounting for a number of changes in the metropolitan area over the decade – including the school finance reform - and then compare the predictions with the 2000 data. The model provides a reasonable fit for the in-sample data used for estimation and the out-of-sample data used for validation, which lends credibility to the policy analysis.

The existing literature on school funding reform encompasses two main types of studies. The first one uses calibrated models to investigate the equilibrium effects of school finance reform.[2] I build on this body of research by estimating an equilibrium model. Since the model lacks a closed-form solution, I apply the one-step, full solution estimator developed in Ferreyra (2007), and enhance it as follows. First, my computational representation of Detroit includes the actual 83 school districts and is thus richer than in previous studies. Second, I model revenues by using the actual state aid formulas rather than the simplifications used in previous estimation,[3] which is critical in light of my interest in school finance reform. Third, I am the first to use school

[2] See, for instance, Nechyba (2004) and Fernandez and Fernandez and Rogerson (2003).
[3] Among other attempts to estimate locational equilibrium models, Calabrese et al (2006) and Ferreyra (2007) use simplifications of actual rules, whereas Bayer et al (2005) do not model revenue determination.

achievement data to identify the importance of peer quality (proxied by parental income) relative to school spending, in contrast with recent studies estimating Tiebout models with peer effects.[4] Although the lack of individual-level data prevents the identification of the mechanisms that give rise to actual peer effects, my estimates indicate that my measure of peer quality is more important than spending in the production of school quality, a finding with relevant implications for policy analysis. In order to use rich structure I overcome a number of empirical and computational challenges. Thus, in this paper I employ elements from the methodological frontier, enhance them, and apply them to a highly relevant and complex policy issue.

Although scholars have recently used their parameter estimates from equilibrium models for policy simulations (Ferreyra 2007), no direct evidence exists on the simulations' plausibility because the simulated policies have never been implemented in large scale. This paper, in contrast, is the first to validate a general equilibrium model of multiple jurisdictions through the simulation of an actual large-scale policy – Proposal A. Researchers have exploited opportunities for model validation that arise due to regime changes, treatment assignment or policy variation.[5] Although useful, this type of exercise is rather uncommon, perhaps because regime shifts are quite rare (Keane and Wolpin 2006). While my framework for large-scale policy analysis is similar to Sieg et al (2004) in the use and estimation of a Tiebout model, it differs in that I allow the equilibrium level of the public good to adjust endogenously in the simulations in contrast to their holding it fixed. Thus, my approach is particularly adept to study large-scale policies.

---

[4] See Calabrese et al (2006) and Ferreyra (2007).

[5] See, for instance, Keane and Moffitt (1998), Keane and Wolpin (2006), Lise et al (2003), Lumsdaine et al (1994), McFadden et al (1997), and Todd and Wolpin (2006).

The second group of studies on school finance reform includes empirical investigations. Most of them are partial equilibrium analyses that have focused on one type of effect,[6] although some have studied general equilibrium effects yet from a reduced-form perspective.[7] In particular, Epple and Ferreyra (2008) investigate the equilibrium effects of Proposal A in the Detroit metropolitan area. While I analyze the effects of Proposal A as well (with results broadly consistent with those of Epple and Ferreyra), I also examine alternative school funding reforms in Michigan. Thus, I am able to quantify each policy's fiscal cost, effects on revenue and school quality, impact on property value and household residential choices, and distributional effects, all of which illuminate the potential political support for each reform. Confidence in these counterfactuals is enhanced by the fact that I estimate and validate the structural model, in contrast to the in-sample, reduced-form investigation from Epple and Ferreyra (2008).

Furthermore, the recent wave of school finance reform litigation has focused on adequacy rather than equity in order to secure for each district the revenue needed for an adequate education (Reich 2006). Hence, an important contribution in this paper is the equilibrium analysis of an adequate funding regime, an exercise not conducted before in the literature. Considering equilibrium effects proves to be crucial in determining adequate revenue. The importance of this paper's counterfactuals cannot be stressed enough, as the sheer dollar amount involved in school funding litigation -particularly over adequacy- points to the need for learning as much as possible about the effects of alternative regimes before incurring the very high costs of implementing any one of them. Moreover, the

[6] See, for instance, Card and Payne (2002), Dee (2000), Downes (1992), Hoxby (2001), Murray et al (1998). For Michigan, see Cullen and Loeb (2004), Guilfoyle (1998), Loeb (2001), Papke (2005), Roy (2003, 2004).
[7] See Keely (2005) for Kentucky, Epple and Ferreyra (2008) and Roy (2004) for Michigan.

framework and insights from this paper can be applied to the analysis of school finance reform in other states, as 18 states are currently undergoing litigation.[8]

My analysis indicates that while Proposal A equalizes revenues to some extent, it is less effective at closing the school quality gap because the reform affects the input with the lesser role, on the margin, in the production of school quality. Thus, the reform only induces small demographic changes, which means that peer quality –and hence school quality- does not change much across districts. The property tax reduction is fully capitalized in housing values, and low-income households are Proposal A's clear gainers.

In addition to Proposal A, I analyze alternative regimes for revenue equity, some of which resemble those recently adopted by other states. In my simulations, even the most effective option -a uniform and high foundation- is quite limited in terms of closing the achievement gap despite its very high fiscal cost. The question then becomes: exactly what is the necessary funding level to secure the desired achievement in each district? Thus, I conduct an adequacy simulation whose main lesson is that when revenue is the only policy lever, even modest increases in achievement are fiscally very costly, and ambitious goals such as the 100 percent proficiency rate targeted by No Child Left Behind are prohibitively costly. These findings point to the need for reforms which do not rely –at least not solely- on revenues. While other studies have reached a similar conclusion,[9] this is the only paper to attain it by using an estimated and validated model, examining a variety of policy reforms, and quantifying each one's equilibrium effects. My study, then, is particularly solid from a theoretical and empirical perspective.

---

[8] The court rulings from November 2006 in *Campaign for Fiscal Equity v. State of New York* exemplify the dollar amounts of this type of litigation. New York's court of appeals ruled that the state must spend an additional $1.93 billion for New York City public schools, short of the almost $5 billion requested by the plaintiffs. For information about current and past school finance litigation, see http://schoolfunding.info.

[9] See the discussion and references in section 7.

The remainder of this paper is organized as follows. Section 2 highlights some changes in the Detroit metropolitan area between 1990 and 2000. Section 3 presents the theoretical model, and Section 4 describes the model's computational version. Section 5 presents the estimation strategy and results, Section 6 the out-of-sample prediction exercise, Section 7 presents the policy analysis, and Section 8 concludes.

2. Detroit in 1990 and 2000

A building block of my framework is the equilibrium computation as a function of the model's exogenous variables. The endogenous variables of interest are district average household income, rental value, spending per student, and school quality. The exogenous variables are the state school finance regime, the metropolitan area income distribution, the district-level stock of non-residential property, and the neighborhood-level quantity and quality of housing. In the estimation (validation) I seek to match the 1990 (2000) values of the endogenous variables given the 1990 (2000) values of the exogenous variables. Hence, in this section I characterize Detroit in 1990 in terms of the endogenous variables, and describe the changes in the exogenous variables over the decade. I focus on the comparison between 1990 and 2000 because demographic and property value data, which come from the Census, are only available every ten years.

Detroit is the largest metropolitan area in the state of Michigan, including eighty-three school districts and a 1990 population of about 3.93 million. In 1990, about a quarter of the population lived in the city of Detroit, which is coterminous with the largest district in the metropolitan area (Detroit Public Schools). Data on income and rental value pertain to households with children in K-12 schools and come from the 1990 and 2000 School District Data Books. Revenues come from the 1989 and 1999 Bulletin 1014 from

Michigan's Department of Treasury. Pass rates for the fourth grade math test, used to measure school quality, come from Michigan's Department of Education.[10] Dollar figures are expressed in 2000 dollars. As Figures 1a and 1b show, in 1990 there was considerable variation in income and housing value across districts, with Detroit Public Schools ranking almost at the bottom. Similarly, local and state revenues differed widely across districts (Figure 1c) as did school achievement (Figure 1d). District average income, rental value, per-pupil revenue, and pass rates were highly and positively correlated.

Proposal A was an important development for the metropolitan area over the decade, and Section 4 describes it in detail. Figure 2 displays revenues the year before the reform ("base revenue") and the foundation allowances guaranteed by the state in 1999. The figure shows that the reform maintained the weak ordering of districts by revenue, and that revenue changes were relatively small in the metropolitan area,[11] although low- and high-revenue districts were clearly the gainers and losers in this reform, respectively. Nonetheless, when measured against revenues in 1989 instead of 1993, the percent gains in revenue over the decade were quite pronounced for several districts (Figure 4a). In addition, in 1991 Michigan implemented the Michigan Educational Assessment Program (MEAP), whose average pass rate rose from 34 percent in 1991 to 70 percent in 2000.[12]

[10] I calculate rental values by annuitizing average owner-occupied house values using the user cost rate. I omit rents in this calculation because the property tax reform applied only to owner-occupied housing units. The series of comparable achievement data begins in 1991. I compute the pass rate as the percent of students who obtain a grade of "satisfactory" in the state's math test. Throughout, demographic data refer to Census years, and school-related data refer to the Fall of the corresponding school year. Revenue, spending and aid are per-student measures. The terms "revenue" and "spending" are interchangeably used.

[11] Revenue changes were more pronounced for rural districts, located outside the Detroit metropolitan area.

[12] The size of these gains must be interpreted with caution, as achievement gains are often quite large when a new test is introduced (Koretz 2002). The expansion of public school choice and the public school accountability implemented in Michigan over the 90s might explain some of the achievement growth, although these programs seem to have gained strength only after 2000 (Cullen and Loeb 2004 and Courant et al 2003) and some still remain quite limited, as documented below. Furthermore, Roy (2003) provides evidence that Michigan's academic gains are much smaller when measured by federal rather than state tests.

Moreover, the metropolitan area income distribution also changed over the decade. While all segments of the income distribution experienced real gains, these were greater for the high and particularly low ends. For instance, at the deciles of the household income distribution on which I focus for computational purposes -10th, 30th, 50th, 70th and 90th percentiles- real income grew by an approximate 24, 9, 8, 12 and 12 percent respectively.

Figure 3 depicts 1990 neighborhood average housing quality.[13] This varied considerably across neighborhoods, with central city neighborhoods being among the lowest. Housing stock in the metropolitan area grew by 6.7 percent, and most of the growth took place in outer suburbs. Of interest in this paper is the change in neighborhood *relative* size, computed as the neighborhood share of the metropolitan area's housing stock (Figure 4b). The central city, in particular, went from 27 to 23 percent of the total stock. Housing qualities also changed (Figure 4c); they improved the most in the outer suburbs although some neighborhoods in the central city improved as well.

To summarize, over the nineties the Detroit metropolitan area experienced changes in aspects regarded as exogenous by the model. My goal is to study whether the model can predict the 1990 and 2000 "snapshots" of Detroit as a function of these aspects.

3. The Model

The model is based on Ferreyra (2007) and Nechyba (1999). In the model, a metropolitan area is populated by a continuum of households, each one endowed with a house. The set of houses in the metropolitan area is partitioned into school districts, and the size of the housing stock equals the measure of endowed houses. Every district d is partitioned into neighborhoods; there are H neighborhoods in total in the metropolitan area. Although

[13] See section 4 for the definition of neighborhood and the calculation of housing quality parameters.

houses may differ in quality across neighborhoods, they have the same quality and rental price within a given neighborhood. The housing stock cannot be varied in quantity or quality. Each household has one child, who must attend a school. Schools are public, and there is one public school in each district. Since a child may only attend the public school where the household resides, choosing locations is equivalent to choosing schools.

Households are heterogeneous in endowment (house plus income, with I income levels in the metropolitan area) and in idiosyncratic preferences for locations. The following Cobb-Douglas utility function describes household preferences:

$$U(\kappa, s, c, \varepsilon) = s^{\alpha} c^{\beta} \kappa^{1-\beta-\alpha} e^{\varepsilon}, \quad \kappa = k_{dh} \qquad (1)$$

where $\alpha, \beta \in (0,1)$, k_{dh} is an exogenous parameter representing the inherent quality of neighborhood h in district d (i.e., housing size and age, geographic amenities, etc.), c is household consumption, s is quality of the child's school, and ε is the household's idiosyncratic preference for the location. For a given household, ε varies across locations. Furthermore, ε is distributed according to a continuous distribution $G(\varepsilon)$, and is independently and identically distributed across locations for a given household and across households.

Household i maximizes utility (1) subject to the following budget constraint:

$$c + (1 + t_d) p_{dh} = (1 - t_y) y_n + p_n \qquad (2)$$

where y_n is the household's income, t_y is the state income tax rate, p_n is the rental price of the household's endowment house, and the right-hand side is the household's total income. Thus, the household chooses to live in location (d, h) with housing price p_{dh} and property tax rate t_d, and uses the remaining income for consumption c.

Schools produce school quality s according to the following production function:

$$s = q^{\rho} x^{1-\rho} \quad (3)$$

where $\rho \in [0,1]$, q stands for the school's average peer quality and x is spending per student at the school. In district *d*, the school's average peer quality is $q_d = \bar{y}_d$, where $\bar{y}_d$ is the average household income in the district. Thus, peer quality captures parental inputs outside spending that are positively associated with household income, such as parental engagement in the student's and the school's activities (McMillan 2000), and parenting skills and home inputs in the production of achievement (Rebell and Wolff 2008, and references therein). District *d*'s spending per student is x_d, funded by a combination of local property and state income taxes as shown below:

$$x_d = t_d \left(P_d + Q_d\right)/n_d + AID_d \quad (4)$$

where n_d is the measure of households in district *d*, AID_d is the state aid per student for district *d*, and P_d and Q_d are the values of residential and non-residential property in the district, respectively.[14] The public school finance regime described in this section applies before the reform; in Section 4 I note the modifications that apply afterwards.

Households choose locations (d, h) and hence schools to maximize their utility subject to their budget constraint, while taking tax rates t_d, district public school qualities s_d, prices $p_{dh,}$, and community compositions as given. Migrating among locations is costless, and a household may choose to live in house other than its endowed house. In addition, households vote on local property tax rates, taking as given their location, property values, the state aid formula explained below, and the choices of others. Households' preferences over property tax rates are single peaked, and property tax rates

[14] For simplicity, I model non-residential property as inelastically supplied and owned by an absentee landlord. Hence, property taxes on non-residential property are fully capitalized, and the gross-of-tax rental price of non-residential property remains constant.

in each district are determined by majority voting.[15]

The state cooperates with district *d* by providing the per-student aid AID_d, funded by a state income tax whose rate t_y balances the state budget constraint. Before Proposal A, the state applies a District Power Equalization (DPE) regime which guarantees a dollar yield per mill[16] levied by guaranteeing a minimum tax base (GTB) *G*, which is exogenously set. Thus, per-student aid for district *d* is given by the following formula:

$$AID_d = \max\left(0, t_d\left(G - \left(P_d + Q_d\right)/n_d\right)\right) \qquad (5)$$

which voters internalize when voting for local property taxes.[17]

An *equilibrium* in this model specifies a partition of the population into districts and neighborhoods, local property tax rates t_d, a state income tax t_y, and house prices p_{dh}, such that: (a) every house is occupied; (b) property tax rates t_d are consistent with majority voting by residents who take their location, property values, and the choices of others as given when voting on local tax rates; (c) the budget balances for each district; (d) the state budget balances, and (e) at prices p_{dh}, households cannot gain utility by moving. Whereas the equilibrium is proved to exist with a finite number of household types (Nechyba 1999), for the case of an infinite number of household types I compute the equilibrium

[15] The current model does not include private schools because the private school sector was not large in Detroit in 1990 (only 11.8 percent of students attended private schools) and it does not seem to have changed much over the decade (less than one percentage point over this period). The presence of private alternatives to public services violates single-peakedness of preferences over tax rates (Stiglitz 1974). Hence, models that include private alternatives (Nechyba 1999, Ferreyra 2007) assume that voters are myopic - they take their choice of public versus private school, their location, property values and others' choices as given when voting on local tax rates. Under this myopia, single-peakedness over property tax rates holds. Voter myopia is the most commonly used assumption for voting behavior even in the absence of private alternatives. See Calabrese et al (2006) and Nechyba (1999) for further discussion and references.

[16] Property tax rates are often expressed in mills. A one-mill rate is equivalent to a rate of 0.1 percent.

[17] The actual formula adds a flat grant, *f*, to the second branch of the max operator. I omit the flat grant because it is very small. Also, voters' internalization of the formula including the flat grant can potentially generate non-single peakedness of preferences over tax rates because for a district with $G<(P+Q)/n$ and a given value of *f*, it is possible to have a positive, negative, or zero value for $f + t(G-(P+Q)/n))$ depending entirely on *t*. Thus, multiple property tax rates may maximize the median voter's utility.

based on equilibrium sufficient conditions.[18]

4. The Computational Version of the Model

The estimation strategy involves computing the equilibrium for the metropolitan area at alternative parameter points to search for the point that minimizes the distance between the predicted equilibrium and the observed 1990 data. The out-of-sample prediction exercise, in turn, involves computing the 2000 equilibrium using the parameter estimates, and comparing it to the observed 2000 data. Since the equilibrium does not have an analytical solution, I solve for it through an iterative algorithm for a tractable representation of the Detroit metropolitan area. I outline below this representation and the algorithm.

Community Structure and Households

My computational representation of the Detroit metropolitan area includes the actual number of districts. I construct neighborhoods such that the central city has the ten neighborhoods identified by the city's actual classification of Census tracts into neighborhoods, and the remaining districts -all of which are relatively very small- have one neighborhood each. A neighborhood's size is proportional to its actual number of housing units in 1990 or 2000 as needed.

In the theoretical model neighborhood h in district d has a neighborhood quality index equal to k_{dh}. Based on Ferreyra (2007), I construct this index for 1990 using 1990

[18] With a finite number of household types, the allocation of households to locations is unique if there is sufficient variation in district average housing quality (Nechyba 1999). Ferreyra (2007) discusses uniqueness of equilibrium in a model with an infinite number of household types. Simulations for a variant of the current model have shown that the equilibrium is robust to the selection of different initial prices and assignments of households to locations, and that the equilibrium at the parameter estimates is locally unique.

Census tract data for the metropolitan area as follows. I regress the logarithm of tract average rental price on housing and neighborhood characteristics and school district fixed effects. Then I compute each tract's neighborhood quality as the tract's fitted rental value net of the school district fixed effect. Thus, this quality measure captures housing and neighborhood characteristics *excluding* school quality. Finally, I set the neighborhood's quality index equal to the quality of the median tract in the neighborhood.

To facilitate the comparison of the 1990 and 2000 neighborhood qualities, I apply the 1990 regression coefficients to the 2000 data to calculate the 2000 fitted rental values and neighborhood quality indexes. This ensures that the 1990 and 2000 indexes for a given tract differ solely because of the observed differences housing and neighborhood characteristics. Since 36 percent of the tracts in the metropolitan area changed boundaries between 1990 and 2000, I use tract-level data from the 1990 Long Form in 2000 Boundaries and the 2000 Long Form, which are normalized to the 2000 boundaries.

As for households, I consider incomes equal to the 10^{th}, 30^{th}, 50^{th}, 70^{th} and 90^{th} percentiles of the income distribution for households with children in K-12 grades in the metropolitan area in 1990 or 2000 as needed. In the computations, income and housing endowments are independently distributed. The equilibrium computation begins with the same income distribution across neighborhoods, equal to the observed metropolitan area's. Since households are assumed to differ in idiosyncratic location preferences, I assume that ε follows a type I extreme value distribution with scale parameter 1/*b* where *b*>0. Thus, $F(\varepsilon) = \exp(-\exp(-\varepsilon / b))$, and the variance of ε equals $(1/6)\pi^2 b^2$.

School Finance

Under the DPE regime prevailing in Michigan until 1993, district *d*'s state aid, AID_d, was determined by the state aid formula in (5), and local property tax revenue by the first term of (4). Although one would expect these expressions to hold when applied to actual data on *P*, *Q*, *n*, *t* and *G*, the fact that they do not means that a model using them would hardly fit the data. Hence, I search for the implicit formula for which (4) and (5) hold.[19] Furthermore, since DPE taxes residential and non-residential property equally but Proposal A does not, I need to quantify each type of property separately to compare policy outcomes. The Appendix describes the implicit formula and the property tax base quantification. Furthermore, under DPE the state funded state aid mostly through income and sales taxes. Since consumption, taxed by the sales tax, is proportional to income, I henceforth simplify by considering only income taxes.[20] Hence, in my computations the state budget constraint is $t_y Y = \sum_d AID_d n_d$, where *Y* is total metropolitan area income.

In contrast with DPE, Proposal A established a foundation grant system by which the state guarantees each district a per-student revenue for operating expenses equal to its foundation allowance. As a function of 1993 base revenues expressed in 1999 dollars (*x* in the formula below), the foundation allowances (*fa*) for 1999 were determined as follows:

---

[19] One reason the DPE formula does not hold when applied to the data is that Michigan uses the DPE formula only to deliver basic (general purpose) aid to school districts, which accounted for almost 60 percent of all state aid in Michigan in 1990/91 (Kearney 1992) whereas the remaining aid was categorical (i.e., intended for specific purposes). Thus, one should not expect the application of the DPE formula to the data to be able to match the total amount of state aid (note that separate data on basic and categorical aid are not publicly available). One would expect, however, that the product of the observed local property tax rate times the observed average state equalized valuation for each district would be (approximately) equal to the observed local revenue per child. This is not the case either, perhaps due to measurement error, despite the fact that the property tax is the only tax that districts may levy to fund public schools. In light of these difficulties, I search for an empirical formula that makes (4) and (5) hold when applied to the actual data.

[20] Epple and Ferreyra (2008) consider the separate role of sales and income taxes. Their main results are unchanged when only income taxes are considered.

$$fa = \begin{cases} 5{,}700 & \text{if } x \le 5{,}502 \\ .833x + 1114.34 & \text{if } x \in (5{,}502,\ 7{,}494) \\ .867x + 860 & \text{if } x \ge 7{,}494 \end{cases} \qquad (6)$$

Under Proposal A, the state requires each district to levy 18 mills on non-residential property and covers the difference between the foundation allowance and this local revenue through a 6-mill tax on residential and non-residential property, and sales and income taxes. Thus, residential property taxes fell from a statewide average of 34 mills to a statewide uniform 6-mill tax. For fiscal reasons the state only guarantees foundation allowances up to the "maximum state guarantee" threshold ($7,200 in 1999). Districts with foundation allowances above this threshold ("hold harmless districts") may levy up to 18 additional mills on residential property to reach their full foundation.

To compute the Proposal A equilibrium, I first compute the DPE equilibrium that would have prevailed in 2000, and then determine foundation allowances by applying (6) to the predicted 2000 DPE revenues. Hence, a district with DPE spending below the maximum state guarantee has a Proposal A spending equal to its foundation allowance fa_d, and a state aid equal to $AID_d = fa_d - \bar{t}_Q Q_d / n_d$, where $\bar{t}_Q$ is the required 18 mills on non-residential property tax rate. Households in this district do not vote for property taxes. In contrast, a hold-harmless district has spending $x_d = \min\left(fa_d, \bar{f} + t_d P_d / n_d \right)$ where $\bar{f}$ is the maximum state guarantee and t_d is the property tax rate, chosen by majority voting, in excess of the 6 mills levied by the state.[21] This district receives state aid equal to $AID_d = \bar{f} - \bar{t}_Q Q_d / n_d$.

[21] This simplifies the actual rule - it does not consider that beyond the first 18 additional mills on residential property it is possible to raise further mills on all property (Adonizio et al 1995). Single- peakednesss may be lost without this simplification. However, very few districts actually raise those additional mills.

Under Proposal A, the state's budget constraint is $t_y Y + \bar{t}\sum_d \left(P_d + Q_d\right) = \sum_d AID_d n_d$,

where Y is total income in the metropolitan area, $\bar{t}$ is the 6 mills levied by the state on all property, AID_d is state aid, and t_y is the income tax rate, whose endogenous value balances the state budget. In all computations voters are subject to the constraint that property tax rates not surpass the 50 mill-maximum permitted by the Michigan constitution.[22]

The Algorithm

In the model, the parameter vector is $\theta = \left(\alpha, \beta, \rho, b\right)$. To compute the equilibrium for a given parameter point, the algorithm iterates as households choose locations and schools and vote for property taxes until no household gains by choosing differently. The input for the algorithm consists of data for the model's exogenous variables –community structure, neighborhood quantity and quality of housing, non-residential property, metropolitan area income distribution, and state aid rule- and the initial distribution of household types and housing prices. The output is the computed equilibrium, which yields the predicted values of the variables of interest in the estimation and out-of-sample prediction.[23]

5. Estimation

In the estimation I match pre-reform, 1990 data for the following district-level variables: y_1=average household income, y_2=average housing rental value, y_3=average spending per student in public schools, and y_4=district fraction of students who pass the fourth grade math test normalized by the metropolitan area's highest fraction (often called "school

---

[22] This millage, which applies to assessed property values, is approximately equal to 200 mills, or 20 percent, when applied to property market values annualized to yield rental values.

[23] More details on the algorithm can be found in Ferreyra (2007).

quality" in what follows).[24] I scale these variables to have unit variance in the sample.

Let D denote the total number of districts in the sample (D=83), and use i for an individual district, with n_i being the number of housing units sampled in district i. Denote by X_i the set of exogenous variables for district i, including all districts' number of neighborhoods, stock of non-residential property, quantity and quality of housing, and data pertaining to the metropolitan area (10th, 30th, 50th, 70th and 90th income percentiles, and school funding regime). I assume the following:

$$E\left(y_{ji} \mid X_i\right) = h_j(X_i, \theta) \qquad j = 1,\ldots 4;\ \ \mathrm{i} = 1,\ldots \mathrm{D} \tag{7}$$

where the h's are implicit nonlinear functions that express the equilibrium value of each endogenous variable I match as a function of the exogenous data and the parameter vector θ. Since the y_{ji}'s are sample means, $C(y_{ji}, y_{ki'} \mid X_i, X_{i'}) = \sigma_{jk} / n_i = \sigma_{jki}$ if i=i' and 0 otherwise, and $V(y_{ji} \mid X_i) = \sigma_{jj} / n_i = \sigma_j^2 / n_i = \sigma_{ji}^2 = \sigma_{jji}$, where σ_{jk} and σ_j^2 denote population covariances and variances, respectively.

I estimate the model using Feasible Generalized Non-Linear Least Squares (FGNLS) and account for heteroskedasticity across observations and cross-equation covariances. In the first stage of FGNLS I find the value of θ that minimizes the following loss function:

$$L(\theta) = \sum_{j=1}^{4} \sum_{i=1}^{D} \left(y_{ij} - \hat{y}_{ij}\left(\theta\right) \right)^2 \tag{8}$$

I use the residuals from this stage to compute the $\hat{\sigma}_{jki}$'s needed to transform the variables

---

[24] The school quality predictions are also normalized by the highest predicted quality in the metropolitan area in order to fit pass rates, which lie between zero and one whether or not they are normalized. The normalization means that the focus is on the achievement gap relative to the highest-achievement district. Hence, an increase in the normalized measure for a given district represents a closing of this gap. In what follows school quality refers to normalized achievement, except when indicated otherwise.

in order to account for heteroskedasticity and cross-equation covariances. In the second stage I minimize the following loss function in the transformed variables:

$$\tilde{L}(\theta) = \sum_{j=1}^{4}\sum_{k=1}^{4}\sum_{i=1}^{D}\left(y_{ji}^{*} - \hat{y}_{ji}^{*}(\theta)\right)\left(y_{ki}^{*} - \hat{y}_{ki}^{*}(\theta)\right) \tag{9}$$

where * denotes division by $\sqrt{\hat{\sigma}_{jki}}$. The value of θ that minimizes this function, $\hat{\theta}$, is the estimate for the parameter vector.

An advantage of estimating the model is the understanding of what features of the data identify each parameter.[25] Spending and housing prices identify the school quality coefficient on the utility function (α), as a higher α raises spending and most housing prices. Housing prices also identify the consumption coefficient in the utility function (β), because a higher β raises consumption and lowers housing prices. The level of spending, and the correlation between income and achievement, identify the elasticity of school quality with respect to peer quality (ρ), as a higher ρ raises this correlation and lowers spending. Finally, the interjurisdictional variation in income, housing prices, spending and achievement identifies *b*, which is directly related to the variance of idiosyncratic preferences, since a greater *b* makes household sorting depend less on income and more on idiosyncratic preferences, and thus leads to less residential segregation across districts.

Table 1 presents the parameter estimates for the model. These are highly significant, mostly as a result of fitting sample means based on large numbers of observations. The estimate of ρ implies that peer quality contributes, on the margin, more

[25] The model is identified if no two distinct parameter points generate the same equilibrium. Formally, a sufficient condition for local identification is that the matrix of first derivatives of the predicted variables with respect to the parameter vector has full column rank when evaluated at the true parameter point. This condition is met when I evaluate that matrix at my parameter estimates.

than spending to achievement. Furthermore, the fact that the estimate for *b* is close to zero implies that households of different incomes do not mix much within districts.

Column 1 of Table 2 shows the root mean squared error for district average income, rental value, spending, and school quality, and Figures 5a through 5d depict the predicted and observed values for these variables. The relatively low root mean squared error for these variables, and the relatively high correlation between predicted and observed values (.84, .88, .76 and .83 for income, rental value, spending and school quality, respectively) indicate a reasonably good fit of the data. This is encouraging given the parsimonious model. The good fit of district average household income and rental value indicates that the model captures locational patterns, although very high income or house values are under predicted because the empirical income distribution is truncated at the 90th percentile. The efforts to quantify property tax bases and to construct implicit funding formulas have helped fit spending. Furthermore, the model fits school quality quite well, with the same caveats noted for income given the high estimate for ρ.

Table 3 shows the correlations between the matched variables for the observed and fitted values. The correlations for fitted values resemble the actual correlations reasonably well. The correlations involving predicted spending are somewhat understated, and the correlations involving predicted school quality are somewhat overstated. However, fitting spending is probably the most challenging aspect of estimation. Furthermore, pass rates measure the corresponding theoretical construct only imperfectly and are likely affected by substantial measurement error (Kane and Staiger 2002). Overall, I view the evidence presented here as indicative that the model successfully captures patterns in the data.

6. Out-of-Sample Prediction

Perhaps more critical for counterfactuals is whether the model can fit out-of-sample data – namely, whether the predicted 2000 equilibrium can replicate the data. If I had the exogenous data needed to compute the 1993/94 equilibrium, I would be able to predict the 2000 foundation allowances because they are a function of 1993/94 revenues. Since such data are not available, to compute the 2000 equilibrium I feed the algorithm with the 2000 value of the exogenous variables and compute the equilibrium that would have prevailed in 2000 had the DPE regime been still operative.[26] I then use (6) to determine foundation allowances given 2000 DPE revenues, and compute Proposal A's equilibrium. Thus, the success of the 2000 predictions partly depends on the counterfactual 2000 DPE.

Pass rates rose consistently over the decade, perhaps due to learning about the test (see Section 2). This improvement displayed a ceiling effect: low-performing districts, with the greatest room for gains, indeed displayed the largest gains. When computing the 2000 equilibrium, I account for this phenomenon by modeling a proportional achievement growth which is larger for low-performing districts and thus consistent with the ceiling effect.[27] As for revenues, a district's predicted 1999 revenue equals the predicted foundation allowance if the district is not allowed to raise hold-harmless mills based on its 2000 DPE revenues. Otherwise, the predicted revenue equals the maximum state guarantee plus the revenue from hold harmless mills.

[26] In order to compute the 2000 DPE equilibrium, I need to choose a value for the GTB. Lacking information on what the GTB would have been in 2000, I choose the observed 1990 value expressed in 2000 dollars. To the extent that this GTB may be too low, the resulting DPE spending for low-revenue districts is also low. Hence, some of the policy effects in section 7 are best viewed as an upper bound.

[27] I do this as follows. I raise all (unnormed) school quality predictions by an additive constant, such that the ratio between this constant and the average predicted 1990 (unnormed) quality is the same as the ratio between the observed average increase in (unnormed) achievement over the decade and the observed 1990 (unnormed) achievement. The adjustment helps match the higher mean and lower variance of achievement in 2000 relative to 1990, which are typical of the implementation of a new test (see Section 2). It is reassuring that predictions with and without the school quality adjustment show very similar patterns. A regression of the school quality predictions including the adjustment on the predictions without the adjustment has an intercept of 0.36, a slope of 0.64, and an R^2 of 0.999. Note that predictions for 2000 without including the school quality adjustment correspond to the Proposal A simulations in next section.

Column 2 of Table 2 depicts the root mean squared error for average income, rental value, foundation allowance, spending and public school quality, and Figures 6a through 6e depict their 2000 predicted and observed values. Overall, the model fits the out-of-sample data reasonably well. There is a bunching of predicted allowances at the minimum foundation ($5,700) because the model under predicts the corresponding DPE revenues. Similarly, the model over predicts DPE revenues (and foundation allowances) for other districts. However, the relatively high correlation between predicted and observed foundation allowances (0.71) reveals an overall good fit for this variable.

Figure 6d displays actual and fitted spending (fitted spending and foundation allowance differ for hold-harmless districts). In the data, 22 districts are allowed to raise hold-harmless mills.[28] The model correctly predicts the hold-harmless status of 14 districts, out of which 8 are predicted to raise hold-harmless mills. Revenue is well fitted overall (the correlation between observed and fitted values equals .71). Figure 6e displays observed and fitted school quality. The model predicts school quality quite well (the correlation between observed and predicted values is .69), although with under prediction for some medium-performance districts and slight over prediction for top districts.

A comparison between panel (a) of Tables 3 and 4 reveals that correlations among the variables of interest did not change much between 1990 and 2000, except that the larger school quality improvement for lower performance districts severed the association between school quality and other variables. As panel (b) of Table 4 shows, the model reasonably replicates the observed 2000 correlations.

---

[28] No direct evidence exists that these districts raised hold-harmless mills in 2000, but the vast majority of these districts did so in 2005/6. See http://www.michigan.gov/mde.

One might ask whether the model captures changes in the endogenous variables occurred over the decade. Hence, Table 5 displays correlations between the observed and predicted changes, the latter computed as the difference between 2000 and 1990 predictions. As the table shows, the model predicts the changes in endogenous variables quite well. Furthermore, the “observed data” row of Table 6 shows the pattern of changes. Low-revenue districts experienced greater absolute and relative increases in revenue, and greater property tax relief. Furthermore, low-income districts experienced greater income growth. One possible explanation is that higher-income households might have migrated towards lower-income locations in response to Proposal A’s incentives. Alternatively, since the low segment of the income distribution experienced the greatest proportional gains in real income (see Section 2), low-income districts might have grown richer simply because their originally inhabitants became richer. I re-examine this matter in Section 7.

Moreover, rental values grew the most in districts with the lowest values, which benefited from the highest revenue increases, largest property tax reductions, and greatest increases in household income. Districts with the lowest initial school quality reaped the largest proportional gains, an outcome likely associated with the ceiling effect. The correlations in the “fitted data” row of Table 6 are encouraging because they show that the model captures the observed pattern of changes.

Table 7 displays correlations among proportional changes in the variables of interest. Correlations among changes in the endogenous variables are reasonably captured by the model, although with some overstatement. Moreover, the model captures the fact that locations with the greatest proportional housing quality increase experienced the greatest increase in household income, rental value and revenue. Nonetheless, actual correlations involving housing quality changes are not particularly high. The correlations involving change in district relative size also seem captured by the model. However, these correlations are almost totally driven by the city of Detroit, because changes in relative size are almost negligible outside the central city, and quite small for the central city. Furthermore, the reported effects of the housing stock reduction in the central city mostly capture effects of the property tax reform. The city's average income, property value, spending and school quality rose by 17, 106, 11 and 178 percent respectively. The model predicts these changes quite well (24, 166, 4 and 254 percent, respectively).

To sum, the model fits the out-of-sample data reasonably well. This lends plausibility to the counterfactual 2000 DPE, and provides confidence for policy analysis.[29]

7. Policy Analysis

In this section I investigate the effects of several school funding reforms including Proposal A. To predict the equilibrium for each funding policy I first compute the benchmark 2000 DPE equilibrium, and then the corresponding policy's equilibrium based on the benchmark. DPE is the natural benchmark because it was the prevailing regime before Proposal A. In order to focus exclusively on funding issues, school quality in these simulations does not incorporate the proportional growth described in Section 6.

Table 8 characterizes the equilibrium for the benchmark DPE and alternative policies by presenting effects on school revenues, demographics, property values, school

[29] The out-of-sample data corresponds to a setting where school districts have little discretion to determine expenditure. This might appear as a limitation for the out-of-sample prediction exercise to the extent that one is interested in counterfactuals that preserve local discretion (as in the DPE simulations examined below). However, the voting model finds some validation with the hold-harmless districts, whose behavior is fit reasonably well. In addition, the model is indirectly validated through the predicted foundation allowances because these are a function of the counterfactual 2000 DPE revenues, which are determined by voting. Furthermore, the voting model fits the in-sample data, where tax rates are determined by voting. Calabrese et al (2006) have also found that the voting model fits the data reasonably well when peer effects are accounted for.

quality and fiscal considerations in panels (a), (b), (c), (d) and (e), respectively. Column 1 pertains to the benchmark DPE equilibrium, in which 14 out of 83 districts have property tax bases per student smaller than the guaranteed tax base (GTB). They receive state aid ($460 on average), funded by income taxes paid mostly by households in high-income districts.[30] Variation (measured by the ratio of the highest to the lowest value and the ratio of the 75th to the 25th percentile) in revenue, income, property values and school quality across districts is considerable. Urban and low-income districts display the lowest income, property values, revenues and achievement in the metropolitan area, and the highest property tax rates. However, fiscal redistribution favors these districts, as the net income tax subsidy per student (per-student aid minus the household's income tax liability) is positive for them yet negative for wealthier districts.

Proposal A

Column 2 of Table 8 displays the effects of Proposal A, and Figure 7 depicts Proposal A revenues relative to the benchmark 2000 DPE. As is clear from (6), the twenty-five districts with 2000 DPE revenue above $6,673 lose funding, whereas the urban and low-income districts with 2000 DPE revenues below $5,502 gain funding. The tax reform favors all districts, although urban and low-income jurisdictions experience the greatest property tax relief because they have the highest property tax rates in the benchmark. Thus, all aspects of Proposal A benefit these districts proportionally the most.

As panel (a) shows, Proposal A reduces the variation in revenue across districts, as it raises urban revenues by 58 percent on average and lowers high-income districts' revenues by 18 percent on average. Furthermore, average income rises slightly in urban

[30] In these simulations, "high-income districts" are those allowed to raise hold-harmless mills in Proposal A.

districts (see panel (b)), because their revenue increase and property tax reduction attracts some higher income households. In contrast, by losing revenues and the ability to choose them, hold-harmless districts lose some high-income households to other districts, particularly those with relatively good housing. As these relocations take place, income variation drops across districts.[31] Nonetheless, the changes in average household income across districts are quite small, which indicates that the reform has little effect on household sorting, a result consistent with Epple and Ferreyra (2008) and Roy (2004). This is because Proposal A affects school spending, which is less important than peer quality in the production of school quality and hence has little ability to affect households' choices. Compounding this problem, housing quality in urban and low-income districts is not high enough to attract many higher income households. Thus, the actual income gains experienced by the lowest income districts over the nineties (see Section 6) are more likely associated with changes in the overall income distribution, which relatively favored original residents in those districts, than with household relocation.

Changes in property values display a similar pattern as changes in income (see panel c). Because they reflect the net effect of lower property taxes and changed school revenues, urban districts attain the largest gains (3 percent on average) and high-revenue districts experience the largest losses (2 percent on average).

An important issue is whether Proposal A affects school quality (see panel (d)). Since peer quality has a prominent role in the production of school quality, and peer qualities do not change much across districts, school qualities change at much lower rates

[31] In the case of income, the highest to lowest ratio is the same as for DPE and the other revenue-equity policies, whereas the 75th to 25th percentile ratio is lower for Proposal A and the other policies than for DPE. Thus, while the relocations favor low-income districts, they are not strong enough to alter the income gap between the highest and lowest income districts.

than revenues. Urban districts gain the most school quality (8 percent on average), while high-income districts lose the most (4 percent on average). Hence, school quality variation shrinks, though not as much as revenue variation. The fact that equalization policies are more effective at equalizing revenues than school quality is a theme in these simulations and shows the limitations faced by state aid policies, a point also made by Nechyba (2004). The contrast between Proposal A's average predicted proportional change in school quality (0.05) and its observed counterpart over the decade (0.77) suggests that little of this increase is associated with Proposal A. Moreover, my results are consistent with others in the literature. In my simulations, a ten percent revenue increase is associated, on average, with about two additional percentage points in the pass rate, similar to estimates reported by Papke (2005) using data for the whole state of Michigan.

Panel (e) reflects the reform's fiscal impact. The average tax burden per household is only slightly higher than in the benchmark. The property tax reform leads to lower residential yet higher non-residential taxes, which are a subsidy to households, and higher income taxes. In contrast with the benchmark, students in high-income districts receive some state aid under Proposal A, although these households also pay greater income tax bills. Hence, their net income tax liability goes from $500 to $4,000. The reverse is true for students in other districts, which go from an average income tax subsidy of $100 to one of $1,100. Even though high-income districts undergo, on average, almost the same tax burden in the benchmark and Proposal A, most of their benchmark burden consists of property taxes to fund their own schools rather than income taxes to fund others' schools.

Even though high-income districts retain some local discretion on tax rates, they still face a revenue cap. Currently 28 states restrict supplementation in some way, and the most radical recent reforms in Kansas, Vermont, Texas and Kentucky also include supplementation limits (Yinger 2004). To investigate the importance of these limits, I simulate a variant of Proposal A that allows these districts to raise their desired level of property taxes. As it turns out, removing the supplementation limits has no effect on high-income districts, which bear most of Proposal A's fiscal cost. In other words, Proposal A affects disposable income in these districts to such an extent that they choose not to raise additional mills even when allowed to do so.[32]

Since Proposal A entails both a tax and a revenue reform, it is interesting to isolate the effects of the tax reform. This reform amounts to adopting a state-wide property tax to fund K-12 schools, as was also the case in California and New Hampshire (McGuire and Papke 2008). Thus, I simulate a reform such that each district's revenue remains the same as in the benchmark, the foundation allowance equals the benchmark revenue, and the tax regime is the same as in Proposal A (the only difference is that hold-harmless districts raise all the residential mills needed to reach their foundation and have the same revenue as in the benchmark). Column 3 of Table 8 displays the effects of this tax reform. Its main impact is on property values, which fully capitalize the reduction in property taxes as gross-of-tax property values are the same, on average, in the benchmark and the tax reform.[33] Hence, the reform leads to overall housing appreciation (4% on average), particularly in urban districts which have the highest benchmark property tax rates. Of all the policies studied here, the tax reform is the only one to increase *all* property values; the other reforms generate gains and losses for low and high-income districts, respectively.

[32] This squares with Yinger (2004), who concludes that reforms that raise state taxes in high-wealth districts will reduce supplementation in those districts to some degree, as seems to have been the case in Kentucky.

[33] If the tax reform replaced residential property taxes only with income taxes, then the gross-of-tax property value *of every house* would be the same before and after the reform. However, the fact that some of the revenue from residential property taxes is replaced by non-residential property taxes creates a subsidy for households, which thus experience a positive yet small income effect. This leads to slight changes in the value of individual houses although the *average* gross of tax property value remains the same.

Furthermore, this full capitalization rate is consistent with the empirical evidence from Epple and Ferreyra (2008) and aggregate-data results from Guilfoyle (1998). My results on tax reform also illuminate the experience in Kentucky, Texas and Vermont, which adopted similar reforms recently (Yinger 2004).

To summarize, one goal of Proposal A was property tax reduction. Another was equity of revenues across districts, and alternative policies could have been implemented to this end. The two main mechanisms commonly used to equalize revenues are District Power Equalization (DPE) and foundation formulas (Yinger 2004), which I analyze below. The benchmark for comparison continues to be the 2000 DPE.

District Power Equalization

As of 1993 Michigan already had a DPE regime aimed at equalizing revenues One reason revenues varied so much across districts in Michigan by 1993 was that policy-makers had allowed the guaranteed property tax base (GTB) to lag behind property values (Cullen and Loeb 2004). Thus, the state could have raised the GTB to secure greater equity. Column 4 of Table 8 shows the effects of doubling the nominal GTB. This amounts to a real GTB increase of about 50 percent, and moves the resulting GTB per student from the initial 30th percentile of the metropolitan area distribution of property tax base per child to slightly above the 40th percentile. This regime ("low GTB") is of interest because the average fiscal burden per household is the same as Proposal A's.

In the low GTB regime, the initial fourteen in-formula districts receive at least twice as much aid as in the benchmark, and ten additional districts are covered by the formula. On average, revenues rise by about 10 percent yet urban districts gain the most. As Figure 7 shows, this policy reduces revenue variation relative to the benchmark although less successfully than Proposal A because revenues depend on local preferences for tax rates. At the same time, the low GTB policy hurts high-income districts less than Proposal A precisely because it preserves local discretion. Since the low GTB policy alters revenues less than Proposal A, demographic and school quality changes are also smaller.

Duncombe and Yinger (1998) discuss the limitations of DPE to achieve revenue equalization. Quantifying the limitations is an empirical matter because of the equilibrium response of voters, property values, and community compositions to DPE incentives. My results suggest that if one remains committed to DPE, greater equalization can only be achieved by setting a very high GTB. Thus, column 5 of Table 8 explores the effects of increasing the initial GTB by a factor of five, which amounts to a real GTB increase of about 270 percent and leaves GTB at the 99th percentile of the distribution of property tax base per student. This regime ("high GTB") amounts to an almost complete equalization of property tax base per student, requiring the same total state aid as Proposal A.

As Figure 7 shows, all districts gain additional revenue relative to the benchmark, and urban districts experience the largest proportional gains. Although revenues are more equally distributed across districts, both relative to the benchmark and to the low GTB regime, they still vary because property tax rates vary. This is consistent with the experience in Missouri reported by Yinger (2004), where GTB was also set at a very high level (the 95th percentile of the distribution of property tax base per student).

While total state aid in the high GTB regime is the same as in Proposal A, total school expenditure is higher because districts can choose their property tax rates. Thus, households undergo a greater property tax burden - greater than in Proposal A, the low GTB regime, and the benchmark. Since the higher revenues do not affect school quality much yet impose a large fiscal burden, property values fall in the vast majority of districts.

This analysis highlights DPE's limitations to achieve revenue equity at a reasonable fiscal cost. Currently, only three states rely exclusively on DPE (Indiana, Missouri and Wisconsin); in these states, revenue differences persist due to differences in local tax rates. Thus, ten states (Florida, Georgia, Iowa, Kansas, Kentucky, Maryland, Montana, Oklahoma, Texas, and Vermont) rely on a two-tiered system that combines foundation and DPE. Large revenue disparity may remain, however, in these combination programs (Picus et al 2008). I now turn to the system that would have completely equalized revenues: a foundation of uniform level across districts.

Uniform Foundation

Basic school finance aid is distributed through a uniform foundation in California and Arkansas. Clearly, the effects of a uniform foundation depend on the foundation level. Whether school finance reform has historically raised or lowered ("leveled up" or "leveled down", respectively) revenues is still an open question.[34] Hence, Columns 6 and 7 of Table 8 display the effects of setting revenues equal to the benchmark median and highest revenue in the "low foundation" and "high foundation" regimes, respectively. For ease of comparison, the tax structure is the same as in Proposal A. Thus, state aid for a district is the difference between the foundation and the district's required property tax revenue. As in Proposal A, the foundation revenue cannot be supplemented.

Figure 7 depicts revenues from the foundation regimes. With the low foundation, urban and suburban districts experience average revenue gains of 69 and 24 percent respectively, yet high-income districts lose at an average rate of 38 percent. Although this

[34] Yinger (2004) and Fischel (2001) review the literature that explores this issue and which yields mixed results. Fischel (2001) examines individual states and concludes that leveling down may have prevailed.

regime induces greater demographic changes than Proposal A, the changes are still small. Moreover, the policy boosts property values in urban and some suburban locations but depresses them in high-income districts.

In contrast to the average revenue gain of 35 percent, the average school quality gain is only 11 percent. School quality variation indeed falls though not as much as revenue variation, showing again the limits of revenue equalization policies. The question, then, is whether equalizing revenue at a higher level would lead to greater school quality gains, a question addressed by the high-foundation simulation. Although the average funding gain for this policy is 228 percent, the demographic effects are the same as for the low foundation. This result, which may be surprising, arises because the two foundation programs eliminate spending as a source of variation across districts, hence leaving housing quality as the only exogenous amenity on which households sort. Since housing qualities are the same in both programs, so are households' choices and school qualities.[35]

Among the policies studied so far, the high foundation leads to the lowest rental value variation because of the large decline in property values in high-income districts, the greatest reliance on income taxes and the highest tax burden. Despite this high cost, the high foundation remains unable to eliminate the achievement gap. Foundations' inability to equalize achievement is consistent with evidence from California's (low) foundation program presented by Downes (1992), who finds virtually no difference in the distribution of achievement across districts before and after California's reform.

Adequacy

[35] Recall that my measure of a district's school quality is the district's achievement normalized by the metropolitan area's highest achievement. According to this measure, school quality is the same for each district in both foundation programs. Absolute (unnormed) achievement, on the other hand, is higher under the high equalization. Relative (normed) achievement highlights the fact that the gap between a given district and the highest achievement district is invariant to the foundation level used for the equalization.

The adequacy movement has sought to secure the funding needed to guarantee a "meaningful" or "adequate" education to all children in a state (Rebell and Wolff 2008), and plaintiffs have adopted a variety of approaches to determine the cost of the desired achievement (Downes and Stiefel 2008). Whereas one can discuss the relative merits of each approach, none of them considers the equilibrium effects of funding changes. For instance, a district's additional funding district might attract higher-income to the district and hence raise its peer quality. Thus, the funding increase needed for an adequate education might be substantially lower than originally thought.

Column 8 of Table 8 presents simulations of an adequacy program that provides revenue so that each district achieves at least 30 percent of the highest district's achievement. For ease of comparison with the other policies examined here, the program awards each district a foundation whose level is the maximum between the district's foundation allowance in Proposal A and the funding required for the achievement target. The tax structure is the same as in Proposal A.

In the simulations, ten districts have a benchmark achievement below the target, and these districts account for about 30 percent of all students in the metropolitan area. Holding peer quality constant, the income tax rate required to finance the adequacy program would be close to 65 percent. However, the revenue increase for the low-performing districts generates some household relocation towards these districts that improves their peer qualities. In equilibrium, the required income tax rate to finance the adequacy program is 10 percent, much lower than the partial equilibrium estimate.

While 30 percent of the highest achievement might not seem an ambitious goal, loftier goals appear to be fiscally very costly. For instance, securing funding so that all districts achieve at least at 40 percent of the highest achieving district entails an equilibrium income tax rate of about 45 percent, and the fiscal burden rises at an increasing rate with the achievement target. The political feasibility of such high fiscal burdens is highly doubtful, which is why I limit my analysis to the goal of having all districts achieve at least at 30 percent of the highest achieving district.

As Figure 7 shows, revenues from the adequacy program are quite similar to those from Proposal A except for a handful of districts, mostly those with initial achievement below the target. These districts attain impressive revenue gains. For instance, the cost of an adequate education in Detroit Public Schools is close to $30,000 per student (as opposed to the $5,700 granted by Proposal A), and between $20,000 and $100,000 for ten other districts. Since state aid is funded mostly though income taxes, high-income districts choose to collect lower property taxes –and obtain lower revenues - than in Proposal A.

As expected, the adequacy program triggers some household relocation. In particular, districts with initial average household income below $20,000 are the chosen destination for some households from districts with average household income between $40,000 and $50,000. These relocations, however, are quite limited relative to the dramatic funding increase received by the lowest achievement districts. Pressed by the higher fiscal burden of this program, some high-income households also leave their original districts in favor of more affordable locations. Except for the poorest districts, which benefit the most through the program, rental values fall everywhere, particularly in high-income districts and even in some of the districts that receive additional funding.

Of all the policies considered here, adequacy is the most effective in terms of school quality. Although the ratio of 75^{th} to 25^{th} school quality percentiles does not change much, highest-to-lowest ratio does. In other words, this policy lifts up the lower tail of the achievement distribution, and brings districts in the upper tail down because they lose revenue and good peers. The counterpart of the achievement gain for the low-performing segment is, of course, the fiscal cost. The average tax burden per household is the highest among the policies considered here, as is the income tax rate. In light of No Child Left Behind's requirement to attain a 100 percent proficiency rate by 2013/14, my simulations suggest that the cost of attaining this goal would be nothing short of prohibitive.[36]

This does not mean, however, that extra resources would always be ineffective; rather, it means that extra resources *per se* might not be very effective. The question, then, is whether other, non-revenue based policies might succeed. Public school accountability is one such policy which seems to have been particularly effective among low performing districts and students (Jacob 2005, Chiang 2007, Rouse et al 2007). As a result, a number of schools have implemented positive behavioral changes (Chiang 2007, Rouse et al 2007, Rebell and Wolff 2008). These effects are noteworthy, particularly because accountability's cost is negligible relative to school funding reform's (Hoxby 2002).

The magnitude of my estimated peer quality parameter suggests that policies aimed at raising peer quality for low-achieving students might be effective. For instance, the state of Michigan implemented open enrollment across school districts in 1996, though the program remains quite small.[37] Perhaps more importantly, my measure of peer quality (household income) is correlated with the quality of parental and home inputs in the

[36] Reich (2006) writes that "…literally to have 'no child left behind' would cost nothing less than the entirety of each state's budget, and even then it is doubtful that the last child would achieve to the adequate standard…". Rebell and Wolff (2008) adhere to this view. Findings from the recent "Getting Down to Facts" initiative in California also echo my results. Imazeki (2006) and Sonstelie (2006) calculate that very large funding increases are needed to meet the desired achievement target in California because poverty has a large effect on achievement yet resources have a modest effect. Even when districts receive extra funding based on adequacy studies, the evidence suggests that they do not use it productively (Picus et al 2008).

[37] Under this legislation, districts choose whether to receive out-of-district applications. According to my calculations, only 5 percent of the K-12 students in the Detroit metropolitan area were participating in these programs in 2006. Not surprisingly, most of the participants resided in low-achievement districts and gained access to better districts through the program. See www.michigan.gov for further detail on the public school choice programs in Michigan.

production of child achievement.[38] The recognition that socioeconomic resources are related to achievement hardly provides a policy prescription to close the achievement gap (Duncan and Magnuson 2005). However, some high-quality interventions engaging children and their parents, particularly before school entry, have proved to be successful (Loeb and Bassok 2008). Furthermore, recent research highlights that the cognitive skills targeted in school reform programs are complementary with non-cognitive skills (such as character traits and personal habits) in the production of human capital (Cunha et al 2006). Thus, the effectiveness of school reform might be further harnessed through educational processes that help students develop both types of skills.

8. Concluding Remarks

In this paper I have presented an empirical framework for large-scale policy analysis. Because of their potential effects on several markets, large-scale policies must be evaluated in an equilibrium framework. Given the high cost of these policies, it is of interest not only to conduct an appropriate evaluation of the observed policies but also to analyze counterfactual policies. I apply my framework to the study of school finance reform in the Detroit metropolitan area, exploiting the occurrence of an actual reform (Proposal A) in 1994. I estimate an equilibrium model of school quality, and household residential and school choice using 1990 data. To validate the model, I use the parameter estimates to predict the 2000 equilibrium, and compare these predictions with 2000 data. The reasonably good fit of the in- and out-of-sample data generates some confidence in

[38] The importance of home inputs is highlighted by Todd and Wolpin (2007), who estimate that about 10 percent of the black-white achievement gap could be closed by equalizing home inputs, and that approximately 50 percent of the gap is attributable to differences in mothers' pre-market skills.

the model for policy analysis. According to my simulations, closing the achievement gap across districts requires a prohibitively large funding increase, even after taking into account mobility effects. These findings are clearly relevant from a policy perspective, given the 100 percent proficiency goal established by No Child Left Behind.

My study relies on a short-run model that assumes that certain variables are exogenous. Thus, it might seem that by specifying the observed value of the exogenous variables, one is in part forcing the direction of the changes between 1990 and 2000 and thus undermining the power of out-of-sample prediction. Perhaps the greatest concern arises over the fixed housing stock, although changes to this stock over the decade do not seem to bear a strong relation with other changes, as noted in the discussion of the out-of-sample exercise.[39] While my attempt to predict the equilibrium at different points in time is no replacement for a long-run, dynamic equilibrium model, I am the first to consider interjurisdictional data from different points in time in an equilibrium framework, and to examine how these points differ. Thus, I view this exercise as an intermediate step between the current empirical Tiebout models, which rely exclusively on cross-sectional data at one point in time and do not conduct model validation (see Bayer et al 2005, Calabrese et al 2006, Ferreyra 2007, and the references therein), and future research that might endogenize the housing stock. The fact that the current model is broadly consistent with the in- and out-of sample data makes it useful for short- and medium-run analysis. Successful out-of-sample exercises increase the confidence in our models and allow us to examine policies of otherwise costly implementation. Any rigor we can bring to this process is certainly desirable, and this paper is a step in that direction.

[39] The concern over changes in the metropolitan area income distribution is lessened by the fact that the changes in the metropolitan area income distribution for Detroit are qualitatively similar to those occurred in the United States as a whole over the nineties (Autor et al 2005).

References

Adonizio, Michael, Philip Kearney and Henry Prince. 1995. "Michigan's High Wire Act". *Journal of Education Finance*, 20 (Winter): 235-269.

Autor, David, Lawrence Katz and Melissa Kearney. 2005. "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists." NBER Working Paper 11627.

Bayer, Patrick, Robert McMillan and Kim Rueben. 2005. "An Equilibrium Model of Sorting in an Urban Housing Market". Duke University Working Paper.

Calabrese, Stephen, Dennis Epple, Thomas Romer and Holger Sieg. 2006 "Local Public Good Provision: Voting, Peer Effects, and Mobility". *Journal of Public Economics*, 90(6-7): 959-981.

Card, David, and Abigail Payne. 2002. "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores". *Journal of Public Economics*, 83(1): 49-82.

Chiang, Hanley. 2007. "How Accountability Pressure on Failing Schools Affects Students Achievement". Working Paper, Harvard University.

Courant, Paul, Julie Berry Cullen, and Susanna Loeb. 2003. **"**K-12 Education in Michigan." In *Michigan at the Millennium: A Benchmark and Analysis Of Its Fiscal and Economic Structure*, ed. Charles Ballard, Paul N. Courant, Douglas C. Drake, Ronald Fischer, and Elisabeth R. Gerber, 1-23. Lansing, MI: Michigan State University Press.

Cullen, Julie Berry and Susanna Loeb. 2004. "School Finance Reform in Michigan: Evaluating Proposal A." In *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. John Yinger, 215:249. Cambridge, MA: MIT Press.

Cunha, Flavio, James Heckman, Lance Lochner, and Dimitryi Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation". In *Handbook of Economics of Education*, ed. Eric Hanushek and Finis Welch.

Dee, Thomas. 2000. "The Capitalization of Education Finance Reforms". *Journal of Law and Economics*, 43 (April): 185-214.

Downes, Thomas. 1992."Evaluating the Impact of School Finance Reform on the Provision of Public Education: The California Case". *National Tax Journal*, 45(4): 405-419.

Downes, Thomas and Stiefel, Leanna. 2008. "Measuring Equity and Adequacy in School Finance." In *Handbook of Research in Education Finance and Policy*, ed. Helen Ladd and Edward Fiske, 222-236. New York: Routledge.

Duncan, Greg and Katherine Magnuson. 2005. "Can Family Socioeconomic Resources Account for Racial and Ethnic Test Score Gaps?". *The Future of Children*, 15(1): 35-54.

Duncombe, William and John Yinger. 1998. "School Finance Reform: Aid Formulas and Equity Objectives". *National Tax Journal*, 51(2): 239-262.

Epple, Dennis and Maria Ferreyra. 2008. "School Finance Reform: Assessing General Equilibrium Effects". *Journal of Public Economics*, 92: 1326-1351.

Fernandez, Raquel and Richard Rogerson. 2003. "Equity and Resources: An Analysis of Education Finance Systems". *Journal of Political Economy*, 111(4): 858-897.

Ferreyra, Maria. 2007**.** "Estimating the Effects of Private School Vouchers in Multi-District Economies". *American Economic Review*, 97 (3): 789-817.

Fischel, William. 2001. *The Homevoter Hypothesis: How Home Values Influence Local Government*

Taxation, School Finance, and Land-Use Policies. Cambridge, MA: Harvard University Press.

Guilfoyle, Jeffrey. 1998. “The Incidence and Housing Market Effects of Michigan’s 1994 School Finance Reforms”. *National Tax Journal Paper and Proceedings.*

Hoxby, Caroline. 2001. “All School Finance Equalizations are Not Created Equal”. *Quarterly Journal of Economics*, 116(4): 1189-1231.

Hoxby, Caroline. 2002. “The Low Cost of Accountability”. NBER Working Paper 8855.

Imazeki, Jennifer. 2006. “Assessing the Costs of K-12 Education in California Public Schools”. In *Getting Down to Facts*, ed. Institute for Research on Education Policy and Practice, Stanford University.

Jacob, Brian. 2005. “Accountability, Incentives and Behavior: Evidence from School Reform in Chicago”. *Journal of Public Economics,* 89(5-6): 761-796.

Kane, Thomas and Douglas Staiger. 2002. “The Promise and Pitfalls of Using Imprecise School Accountability Measures”. *Journal of Economic Perspectives*, 16(4): 91-114.

Keane, Michael and Robert Moffitt. 1998. “A Structural Model of Multiple Welfare Program Participation and Labor Supply”. *International Economic Review,* 39(3): 553-589.

Keane, Michael and Kenneth Wolpin. 2006.“Exploring the Usefulness of a Non-Random Holdout Sample for Model Validation: Welfare Effects on Female Behavior”. Penn Institute for Economic Research Working Paper 06-006.

Kearney, C. Philip. 1992. “Michigan”. In *Public School Finance Programs of the United States and Canada 1990-91,* ed. Steven Gold et al, 287-299. Albany: American Education Finance Association.

Keely, Louise. 2005. The Myth of School Finance Equalization, as told by Kentucky and North Carolina. University of Wisconsin-Madison Working Paper.

Koretz, Daniel. 2002. “Limitations in the Use of Achievement Tests as Measures of Educators’ Productivity”. *Journal of Human Resources,* 37(4): 752-777.

Lise, Jeremy, Shannon Seitz, and Jeffrey Smith. 2003. “Equilibrium Policy Experiments and the Evaluation of Social Programs”. Queen’s University Working Paper 1012.

Loeb, Susanna. 2001. “Estimating the Effects of a School Finance Reform: A Framework for a Federalist System”. *Journal of Public Economics*, 80: 225-247.

Loeb, Susanna and Daphna Bassok. 2008. “Early Childhood and the Achievement Gap”. In *Handbook of Research in Education Finance and Policy,* ed. Helen Ladd and Edward Fiske, 517-534. New York: Routledge.

Lumsdaine, Robin, James Stock, and David Wise. 1994. “Pension Plan Provisions and Retirement: Men and Women, Medicare, and Models”. In *Studies in the Economics of Aging*, ed. David Wise, 183-212. Chicago, IL: University of Chicago Press.

McFadden, Daniel, and Antti P. Talvitie and Associates. 1997. “Validation of Disaggregate Travel Demand Models: Some Tests”. In *Urban Demand Forecasting Project, Final Report: Volume V: Demand Model Estimation and Validation.* Part II, Chapter 3. Berkeley: Institute of Transportation Studies, University of California, Berkeley.

McGuire, Therese and Leslie Papke. 2008. “Local Funding of Schools: The Property Tax and Its Alternatives”. In *Handbook of Research in Education Finance and Policy*, ed. Helen Ladd and Edward Fiske, 357-372. New York: Routledge.

McMillan, Robert. 2000. “Parental Involvement and Competition: An Empirical Analysis of Public School Quality”. University of Toronto Working paper.

Murray, Sheila, William Evans, and Robert Schwab. 1998. “Education Finance Reform and the Distribution of Education Resources”. *American Economic Review,* 88: 789-812.

Nechyba, Thomas. 2004. "Prospects for Achieving Equity or Adequacy in Education: The Limits of State Aid in General Equilibrium". In *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. John Yinger, 111-143. Cambridge, MA: MIT Press.

Nechyba, Thomas. 1999. “School Financed Induced Migration Patterns: The Case of Private School Vouchers”. *Journal of Public Economic Theory*, 1(1): 5-50.

Papke, Leslie. 2005. The Effects of Spending on Test Pass Rates: Evidence From Michigan. *Journal of Public Economics*, 89(5-6): 821-839.

Picus, Lawrence, Margaret Goertz and Allan Odden. 2008. “Intergovernmental Aid Formulas and Case Studies”. In *Handbook of Research in Education Finance and Policy*, ed. Helen Ladd and Edward Fiske, 257-274. New York: Routledge.

Rebell, Michael and Jessica Wolff. 2008. *Moving Every Child Ahead.* New York: Teachers College Press.

Reich, Rob. 2006. “Equality and Adequacy in the State’s Provision of Education: Mapping the Conceptual Landscape.” In *Getting Down to Facts*, ed. Institute for Research on Education Policy and Practice, Stanford University.

Rouse, Cecilia, Jane Hannaway, Dan Goldhaber, and David Figlio. 2007. “Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure”. NBER Working Paper 13681.

Roy, Joydep. 2003. “Impact of School Finance Reform on Resource Equalization and Academic Performance: Evidence From Michigan.” Economic Policy Institute University Working Paper.

Roy, Joydeep. 2004. “Effect of a School Finance Reform on Housing Variables and Socieconomic Segregation: Evidence from Michigan.” Economic Policy Institute University Working Paper.

Sieg, Holger, Kerry Smith, Spencer Banzhaf, and Randall Walsh. 2004 “Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods”. *International Economic Review*, 45(4): 1047-1077.

Sonstelie, Jon. 2006. “Aligning School Finance with Academic Standards”. In *Getting Down to Facts*, ed. Institute for Research on Education Policy and Practice, Stanford University.

Stiglitz, Joseph. 1974**. “**The demand for Education in Public and Private School Systems”. *Journal of Public Economics*, 44: 217-237.

Todd, Petra and Kenneth I. Wolpin. 2006. “Assessing the Impact of a School Subsidiy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility”. *American Economic Review*, 96(5): 1384-1417.

Todd, Petra and Kenneth I. Wolpin. 2007. “The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps”. *Journal of Human Capital*, 1(1): 91-136.

Yinger, John. 2004. “State Aid and the Pursuit of Educational Equity: An Overview.” In *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*, ed. John Yinger, 3-57. Cambridge, MA: MIT Press.

TABLE 1
Parameter Estimates

Parameter	Estimates
α	0.137 (0.003)
β	0.740 (0.003)
ρ	0.871 (0.002)
b	0.006 (0.001)
Sum of Squared Residuals	569.669

Standard Errors in parentheses. Number of observations: 83 school districts.

TABLE 2
Root Mean Squared Error for In- and Out-Of-Sample Predictions

	In-Sample (1)	Out-Of-Sample (2)
Income	1.89	2.27
	(6.44)	(7.39)
Rental Value	0.40	0.60
	(1.26)	(2.02)
Spending	0.17	0.14
	(0.64)	(0.75)
School Quality	0.15	0.15
	(0.48)	(0.77)

Number of observations: 83 districts. Data refer to district averages, and are weighted by district number of housing units. Sample means are in parentheses. Income, rental value and spending are expressed in $10,000. School quality ranges between 0 and 1.

TABLE 3
In-Sample Goodness of Fit: Some Correlations

a. Observed Data

	Income	Rental Value	Spending	School Quality
Income	1			
Rental Value	0.99	1		
Spending	0.66	0.65	1	
School Quality	0.85	0.86	0.59	1

b. Fitted Data

	Income	Rental Value	Spending	School Quality
Income	1			
Rental Value	0.98	1		
Spending	0.46	0.44	1	
School Quality	0.99	0.98	0.54	1

Number of observations: 83 districts. Data refer to district averages, and correlations are weighted by district number of housing units.

TABLE 4
Out-of-Sample Goodness of Fit: Some Correlations

a. Observed Data

	Income	Rental Value	Foundation Allowance	School Quality
Income	1			
Rental Value	0.98	1		
Foundation Allowance	0.63	0.61	1	
School Quality	0.69	0.70	0.46	1

b. Fitted Data

	Income	Rental Value	Foundation Allowance	Spending	School Quality
Income	1				
Rental Value	0.99	1			
Foundation Allowance	0.38	0.38	1		
Spending	0.41	0.41	0.97	1	
School Quality	0.99	0.99	0.43	0.45	1

Number of observations: 83 districts. Data refer to district averages, and correlations are weighted by district number of housing units.

TABLE 5
Out-Of-Sample Goodness of Fit: Correlation between Observed and Predicted Changes

Income	Rental Value	Revenue	School Quality
0.39	0.46	0.60	0.59

Number of observations: 83 districts. Data refer to the correlation between district predicted and observed change in the corresponding average.

TABLE 6
Out-Of-Sample Goodness of Fit: Correlations Between 1990 Values and Percent Changes

	Income	Rental Value	Revenue	School Quality
Observed Data	-0.23	-0.63	-0.72	-0.83
Fitted Data	-0.26	-0.72	-0.76	-0.79

Number of observations: 83 districts. Data refer to district averages and percent change in district averages. For instance, -0.23 under "Income" for the observed data means that the correlation between 1990 observed income and the percent income change over the decade is -0.23, and -0.26 under "Income" for the fitted data means that the correlation between 1990 predicted income and the predicted percent change is -0.26.

TABLE 7
Out-of-Sample Goodness of Fit: Correlations between Changes

a. Observed Data

	Income	Rental Value	Spending	School Quality	Housing Quality	Relative Size
Income	1					
Rental Value	0.49	1				
Spending	0.13	0.27	1			
School Quality	0.22	0.59	0.20	1		
Housing Quality	0.35	0.20	0.26	0.05	1	
Relative Size	-0.35	-0.72	-0.03	-0.73	0.05	1

b. Fitted Data

	Income	Rental Value	Spending	School Quality	Housing Quality	Relative Size
Income	1					
Rental Value	0.75	1				
Spending	0.47	0.40	1			
School Quality	0.67	0.96	0.50	1		
Housing Quality	0.48	0.25	0.27	0.09	1	
Relative Size	-0.17	-0.67	0.21	-0.56	0.05	1

Number of observations: 83 districts. Data refer to percent change in district average for income, rental value, spending, school quality and housing quality, and change in district relative size. Correlations are weighted by district number of housing units.

TABLE 8
Effects of Alternative Policies

a. Revenue per Student

	DPE (1)	Proposal A (2)	Tax Reform (3)	Low GTB (4)	High GTB (5)	Low Foundation (6)	High Foundation (7)	Adequacy (8)
Average Revenue	$5,700	$6,400	$5,700	$6,100	$9,900	$6,100	$14,900	$19,200
Urban Districts	$3,600	$5,700	$3,600	$4,500	$8,800	$6,100	$14,900	$42,200
Suburban Districts	$6,400	$6,700	$6,400	$6,600	$10,200	$6,100	$14,900	$12,100
High-income Districts	$10,100	$8,200	$10,100	$10,100	$13,000	$6,100	$14,900	$8,600
Highest / Lowest	10.06	2.30	10.06	7.51	4.22	1.00	1.00	99.59
75^{th} pctile. / 25^{th} pctile.	2.21	1.26	2.21	1.84	1.28	1.00	1.00	1.26
Avg. Proportional Change		0.34	0.00	0.10	0.96	0.35	2.28	5.10
Urban Districts		0.58	0.00	0.26	1.43	0.69	3.14	11.30
Suburban Districts		0.26	0.00	0.05	0.81	0.24	2.02	3.17
High-income Districts		-0.18	0.00	0.00	0.32	-0.38	0.52	-0.12

b. Household Income

	DPE (1)	Proposal A (2)	Tax Reform (3)	Low GTB (4)	High GTB (5)	Low Foundation (6)	High Foundation (7)	Adequacy (8)
Average Income	$66,700	$66,700	$66,700	$66,700	$66,700	$66,700	$66,700	$66,700
Urban Districts	$26,700	$26,900	$26,900	$26,700	$25,700	$27,100	$27,000	$27,000
Suburban Districts	$79,100	$79,100	$79,100	$79,100	$79,400	$79,000	$79,000	$79,000
High-income Districts	$94,100	$91,000	$93,800	$93,400	$91,000	$88,300	$88,300	$89,200
Highest / Lowest	8.40	8.40	8.40	8.40	8.40	8.40	8.40	7.62
75^{th} pctile. / 25^{th} pctile.	2.54	2.42	2.53	2.50	2.49	2.34	2.34	2.37
Avg. Proportional Change		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Urban Districts		0.01	0.01	0.00	-0.04	0.01	0.01	0.01
Suburban Districts		0.00	0.00	0.00	0.03	0.00	0.00	0.03
High-income Districts		-0.04	0.00	-0.01	-0.04	-0.08	-0.08	-0.07

c. Rental Value

	DPE (1)	Proposal A (2)	Tax Reform (3)	Low GTB (4)	High GTB (5)	Low Foundation (6)	High Foundation (7)	Adequacy (8)
Average Rental Value	$21,500	$21,900	$22,300	$21,300	$20,900	$21,800	$21,100	$18,300
Urban Districts	$12,100	$12,400	$12,600	$12,100	$11,900	$12,400	$12,300	$10,600
Suburban Districts	$24,400	$24,800	$25,300	$24,200	$23,700	$24,700	$23,900	$20,700
High-income Districts	$28,200	$27,800	$28,700	$27,800	$26,500	$26,900	$26,000	$23,000
Highest / Lowest	4.49	4.48	4.57	4.46	4.32	4.45	4.27	3.90
75^{th} pctile. / 25^{th} pctile.	2.05	1.97	2.00	2.02	1.98	1.91	1.88	2.12
Avg. Proportional Change		0.02	0.04	-0.01	-0.02	0.02	-0.01	-0.15
Urban Districts		0.03	0.04	0.00	-0.02	0.03	0.02	-0.12
Suburban Districts		0.02	0.03	-0.01	-0.02	0.01	-0.02	-0.15
High-income Districts		-0.02	0.02	0.00	-0.06	-0.05	-0.08	-0.20

d. School Quality

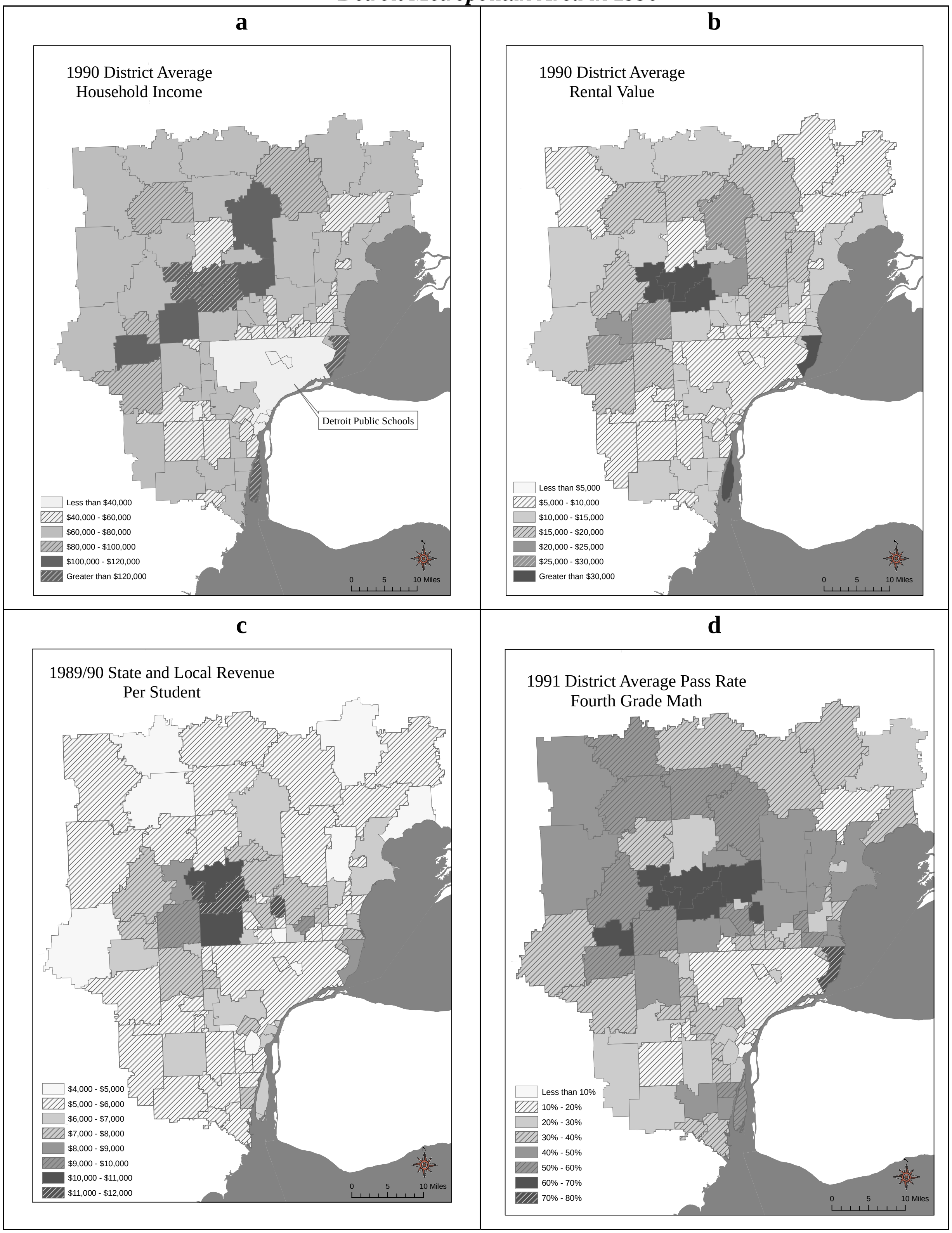
	DPE (1)	Proposal A (2)	Tax Reform (3)	Low GTB (4)	High GTB (5)	Low Foundation (6)	High Foundation (7)	Adequacy (8)
Average School Quality	0.50	0.52	0.51	0.51	0.52	0.54	0.54	0.55
Urban Districts	0.22	0.23	0.22	0.22	0.23	0.25	0.25	0.30
Suburban Districts	0.59	0.60	0.60	0.60	0.61	0.63	0.63	0.63
High-income Districts	0.73	0.70	0.73	0.73	0.72	0.70	0.70	0.71
Highest / Lowest	8.21	6.81	8.20	7.89	7.40	6.38	6.38	3.33
75^{th} pctile. / 25^{th} pctile.	2.27	2.22	2.26	2.25	2.22	2.10	2.10	2.17
Avg. Proportional Change		0.05	0.00	0.01	0.06	0.11	0.11	0.22
Urban Districts		0.08	0.01	0.03	0.05	0.16	0.16	0.41
Suburban Districts		0.04	0.00	0.01	0.07	0.09	0.09	0.16
High-income Districts		-0.04	0.00	-0.01	-0.03	-0.06	-0.06	-0.05

e. Fiscal Implications

	DPE (1)	Proposal A (2)	Tax Reform (3)	Low GTB (4)	High GTB (5)	Low Foundation (5)	High Foundation (7)	Adequacy (8)
Residential Property Tax Rate								
Average	0.07	0.02	0.03	0.07	0.07	0.02	0.02	0.02
Minimum	0.04	0.02	0.02	0.04	0.04	0.02	0.02	0.02
Maximum	0.09	0.06	0.10	0.09	0.09	0.02	0.02	0.02
Non-Resid. Property Tax Rate								
Average	0.07	0.10	0.10	0.07	0.07	0.10	0.10	0.10
Minimum	0.04	0.10	0.10	0.04	0.04	0.10	0.10	0.10
Maximum	0.09	0.10	0.10	0.08	0.09	0.10	0.10	0.10
Income Tax Rate	0.00	0.02	0.01	0.00	0.03	0.02	0.07	0.10
Avg. Tax Burden per Household	$1,700	$1,800	$1,400	$1,800	$3,300	$1,600	$5,000	$6,800
Avg. Net Income Tax Subsidy per Student	$0	$0	$0	$0	$0	$0	$0	$0
High-income Districts	-$500	-$4,000	-$2,600	-$1,100	-$4,500	-$5,100	-$11,100	-$25,700
Low-income Districts	$100	$1,100	$700	$300	$1,200	$1,400	$3,100	$6,600
Avg. Share of Residential Property Taxes	0.69	0.21	0.30	0.64	0.38	0.22	0.09	0.15
Avg. Share of Non-Resid. Prop Taxes	0.22	0.26	0.31	0.21	0.13	0.30	0.12	0.21
Avg. Share of Income Taxes	0.09	0.53	0.39	0.15	0.49	0.48	0.79	0.64

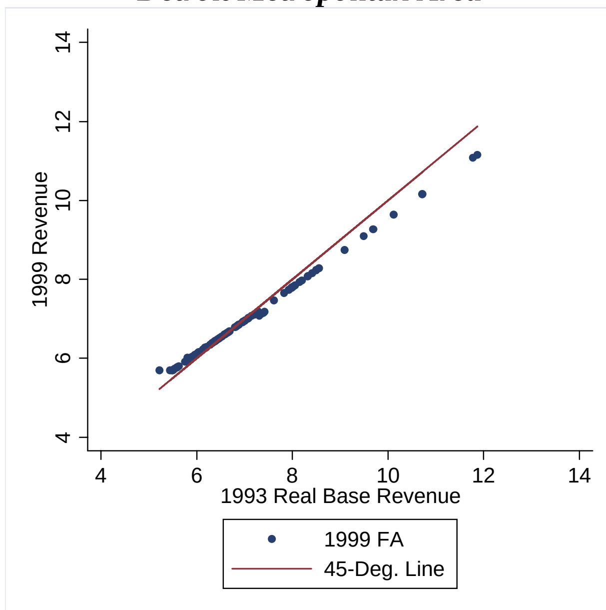
Number of observations: 83 districts. For a given variable, "Highest/Lowest" is the ratio of the metropolitan area's highest to lowest district average, "75th pctile. /25th pctile," is the ratio of the metropolitan area's 75th to 25th percentile, and "average proportional change" is the average of the proportional change in district averages. Changes are computed relative to the benchmark 2000 DPE. Dollar figures rounded to closest hundred. All averages are weighted; weight is number of housing units.

FIGURE 1
Detroit Metropolitan Area in 1990



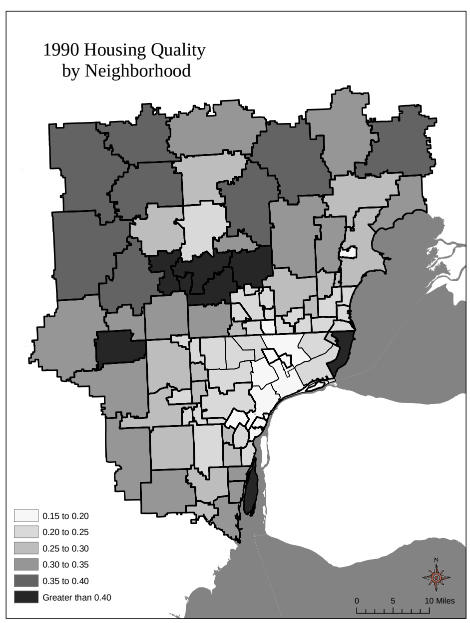
Source: 1990 School District Data Book, Michigan Department of Treasury, and Michigan Department of Education.

FIGURE 2
1999 Real Foundation Allowance and 1993 Base Revenue
Detroit Metropolitan Area



FA: Foundation Allowance. Revenue and foundation allowance expressed in thousands of 2000 dollars.
Source: Michigan Bulletin 1014.

FIGURE 3
Districts, Neighborhoods and 1990 Housing Quality in the Detroit Metropolitan Area



The thicker black lines represent school district boundaries, and the thin lines represent neighborhood boundaries.
Source: author's calculations using 1990 Census data.

FIGURE 4
Changes in the Detroit Metropolitan Area

Share of housing stock is expressed as a percent relative to the metropolitan area, and change in share is expressed in percentage points. Source: Michigan Department of Treasury, School District Data Book, and author's own calculations based on 1990 and 2000 Census.

FIGURE 5
In-Sample Goodness of Fit: Fitted vs. Observed Values

Figure 5a - Average Household Income (in $10,000)

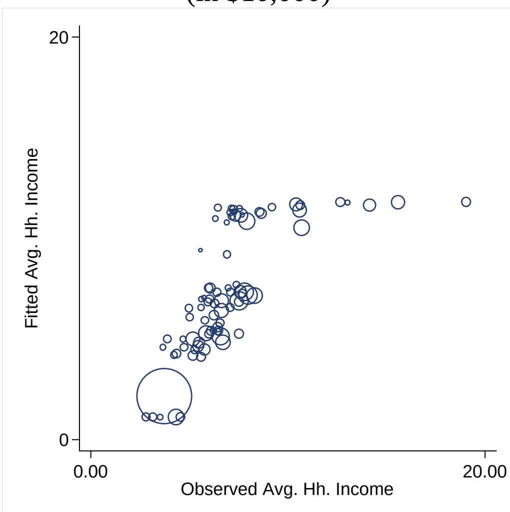


Figure 5b - Average Rental Value (in $10,000)

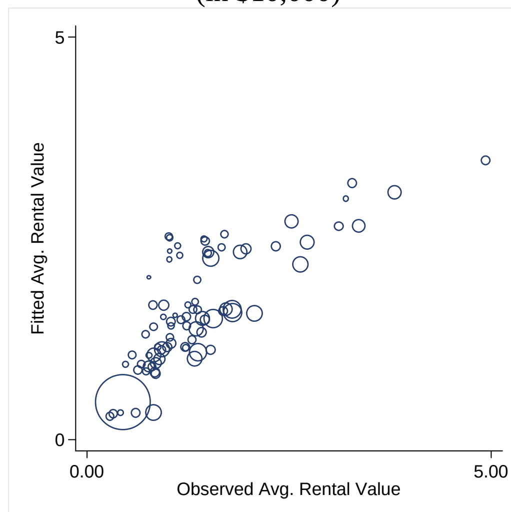


Figure 5c – Spending per Student in Public Schools (in $10,000)

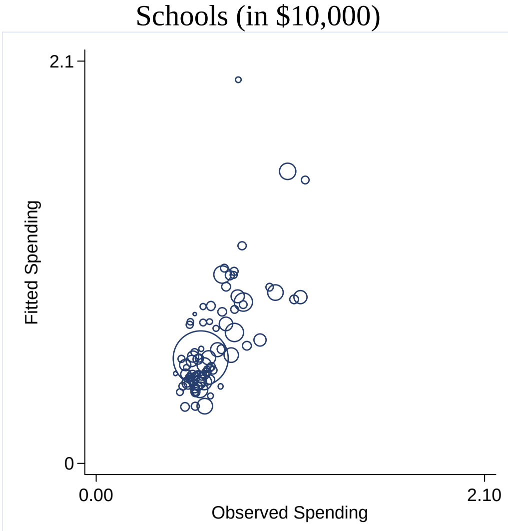
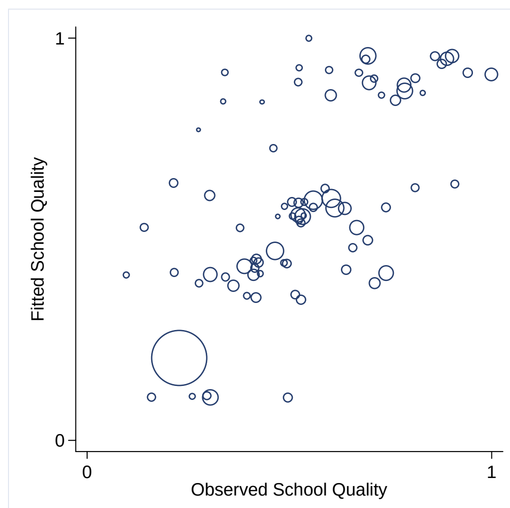


Figure 5d – Public School Quality



Note: observed values on the horizontal axis; fitted values on the vertical axis. Circle size is proportional to the observation's total measure of households. Correlations between fitted and observed values are weighted by the observations' measure of households and are as follows: 0.84, 0.88, 0.76, and 0.83 for Figs. 5a through 5d respectively.

FIGURE 6 – *Out-Of-Sample Goodness of Fit: Fitted vs. Observed Values*

Figure 6a – Average Household Income
(in $10,000)

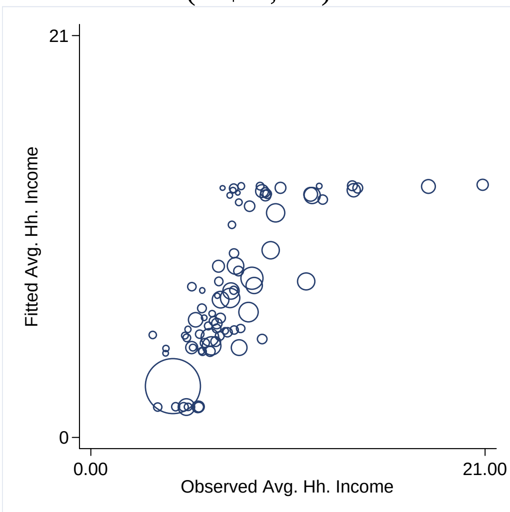


Figure 6b - Average Rental Value
(in $10,000)

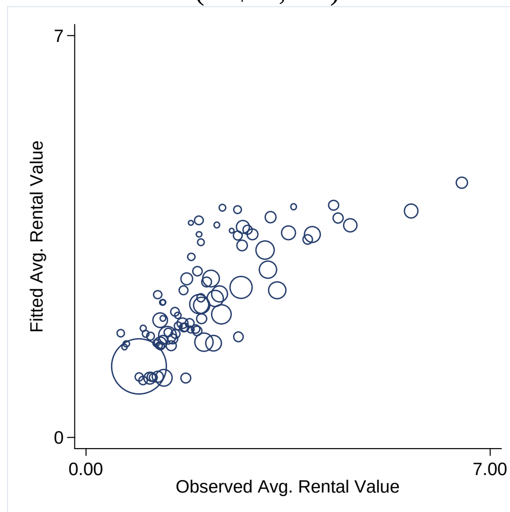


Figure 6c – Foundation Allowance
(in $10,000)

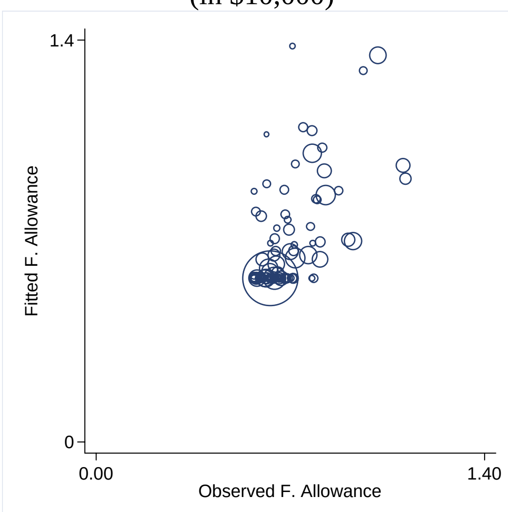


Figure 6d – Per-Pupil Spending
(in $10,000)

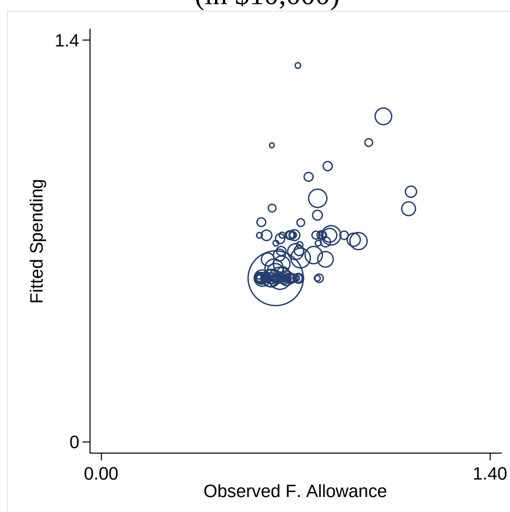
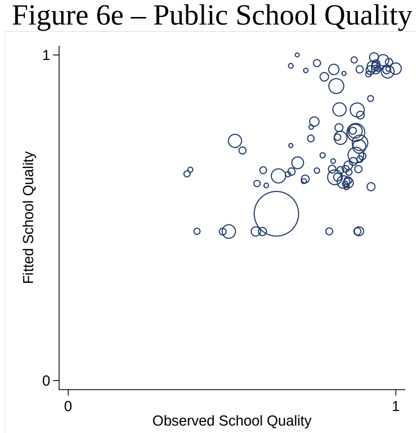
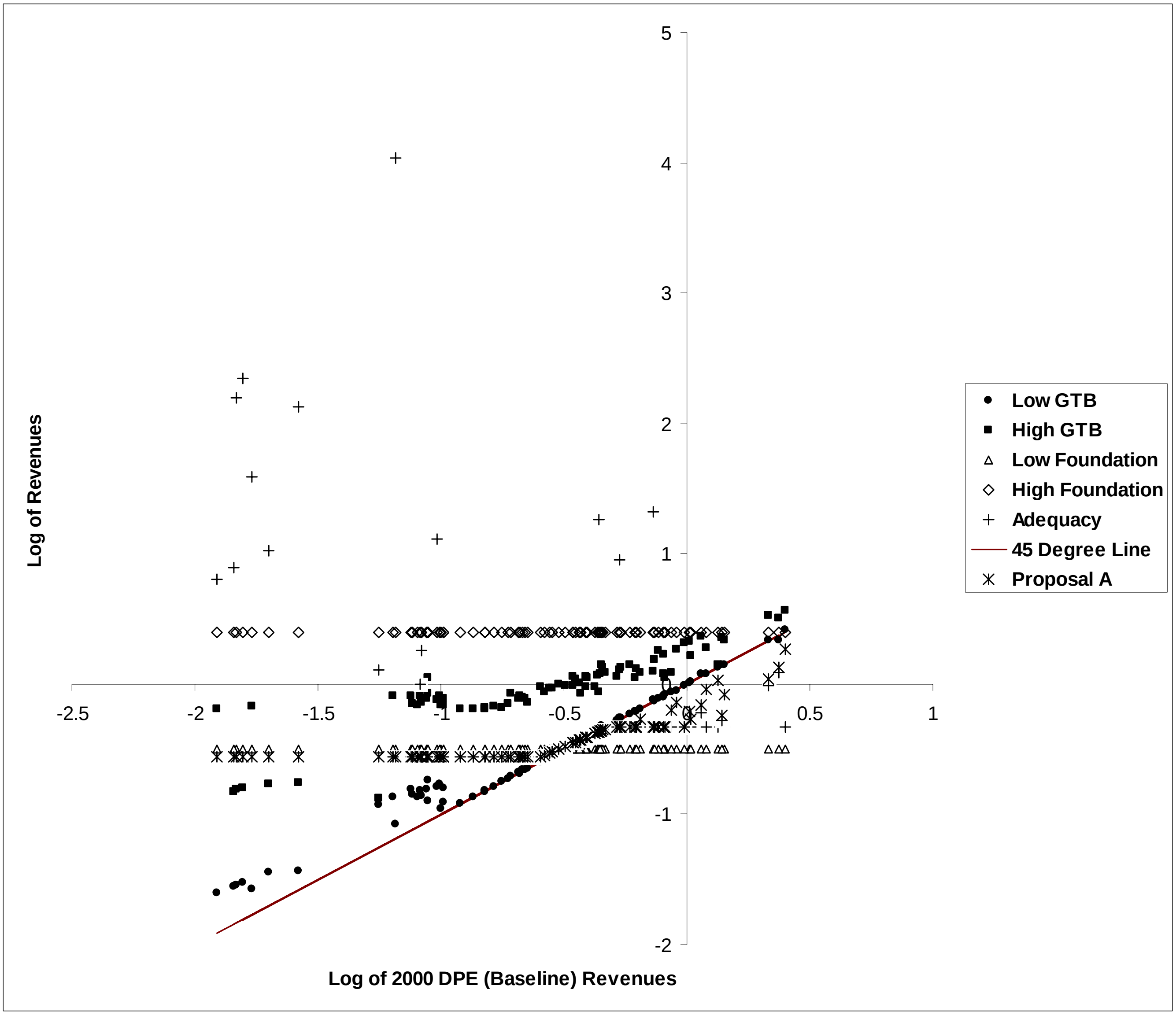


Figure 6e – Public School Quality



Note: see Figure 5. Correlations between fitted and observed values are weighted by the observations' measure of households and are as follows: .82, .86, .71 .71, .69 for Figs. 6a through 6e respectively.

FIGURE 7
Log of Per-Student Revenue under Alternative Regimes



Note: "log" stands for natural logarithm. Revenues are expressed in $10,000. Source: author's simulations.

Appendix

Under the DPE regime prevailing in Michigan until 1993, district *d*'s state aid, A_d, was determined by the following state aid formula:

$$A_d = \max\left(0, f + \hat{t}_d\left(\hat{G} - E_d\right)\right) \qquad \text{(A1)}$$

where f is a flat grant, $\hat{t}_d$ is the district's millage raised for operational purposes, $\hat{G}$ is the guaranteed level for the assessed tax base per student, and E_d is the district's per-student assessed valuation of the property tax base. In Michigan, property is assessed at half of its market value, and in 1989 the nominal values of f and $\hat{G}$ were \$310 and \$83,610, respectively.

In principle, one would expect (1) to hold when applied to actual data on *AID*, $\hat{t}$, and *E*. One would also expect the following expression for local property tax revenue in district *d*, l_d, to hold when applied to actual data:

$$l_d = \hat{t}_d E_d \qquad \text{(A2)}$$

Since neither (A1) nor (A2) hold for actual data,[40] I search for the implicit formulas for which (A1) and (A2) hold when applied to actual data. According to these formulas,

$$AID_d = \max\left(0, \hat{t}_d (G - \phi_d \overline{P}_d)\right) \qquad \text{(A3)}$$

$$l_d = \hat{t}_d \gamma_d \overline{P}_d \qquad \text{(A4)}$$

where $\overline{P}_d$ is half of the observed average house value in district *d*.[41]

To compare DPE and Proposal A outcomes in policy simulations I need to quantify

[40] The expression in (A1) does not hold in the data whether or not the flat grant *f* is included in the formula.
[41] I use half of the observed average house value because the observed tax rates apply to assessed values.

residential and non-residential property separately, because DPE treats both types of property equally but Proposal A does not. Furthermore, this quantification needs to be consistent with the implicit formulas (A3) and (A4). Hence, for 1990 I proceed as follows. I express the per-student total property tax base in district *d* as $\bar{P}_d h_d \left(1 + r_d\right)$, where h_d is the number of households per child in the district and r_d is the ratio of non-residential to residential property. This recognizes the fact that the observed average house value in any given district differs from the per-student property tax base because not every household has one child, and because the property tax base includes non-residential besides residential property. This allows me to express the per-student residential property tax base as $\bar{P}_d h_d$ and the per-student non-residential property tax base as $\bar{P}_d h_d r_d$. Consistency with the implicit formulas requires a relationship among ϕ_d, γ_d, h_d and r_d for each district. Hence, I run the following regressions using 1990 data:[42]

$$\hat{\phi}_d = a + b_1 h_d + b_2 h_d r_d \qquad \text{(A5)}$$

$$\hat{\gamma}_d = \tilde{a} + \tilde{b}_1 h_d + \tilde{b}_2 h_d r_d \qquad \text{(A6)}$$

I then quantify the per-student residential property tax base as $P_d^a = \tilde{\phi}_{Pd}^a \bar{P}_d = \frac{\phi_d}{\hat{\phi}_d}\left(\frac{a}{2} + b_1 h_d\right)\bar{P}_d$ and the per-student non-residential property tax base as $Q_d^a = \tilde{\phi}_{Qd}^a \bar{P}_d = \frac{\phi_d}{\hat{\phi}_d}\left(\frac{a}{2} + b_2 h_d r_d\right)\bar{P}_d$ to fit the aid implicit formula (A3). Similarly, I quantify the per-student residential and non-residential property tax base, respectively, as $P_d^x = \tilde{\gamma}_{Pd}^x \bar{P}_d = \frac{\gamma_d}{\hat{\gamma}_d}\left(\frac{\tilde{a}}{2} + \tilde{b}_1 h_d\right)\bar{P}_d$ and $Q_d^x = \tilde{\gamma}_{Qd}^x \bar{P}_d = \frac{\gamma_d}{\hat{\gamma}_d}\left(\frac{\tilde{a}}{2} + \tilde{b}_2 h_d r_d\right)\bar{P}_d$ to fit the spending implicit formula (A4). It can be verified that

[42] Data on *h* come from the School District Data Book, and data on *r* come from the Citizens' Research Council of Michigan.

$P_d^a + Q_d^a = \phi_d \bar{P}_d$, which is the value of the property tax base for which (A3) holds, and $P_d^x + Q_d^x = \gamma_d \bar{P}_d$, which is the value of the property tax base for which (A4) holds.

In (A1) through (A4), property values refer to assessed values, and tax rates are the millages applied to assessed values. In contrast, the model and computational applications use market rental prices for property values, and the millages are scaled correspondingly. The relationship between a given $\hat{t}$ which applies to assessed property values and *t*, which applies to rental values, is $t = \left(\frac{1}{2u}\right)\hat{t}$, where *u* is the user cost on owner-occupied housing units, equal to 0.12 for the 1990.[43] Hence, in the computational version of the model $AID_d = \max\left(0, t_d\left(G - (P_d^a + Q_d^a)\right)\right)$ and $x_d = t_d\left(P_d^x + Q_d^x\right) + AID_d$, where G is the GTB, equal to $20,066 (=$83,610*0.12*2, given that it applies to property rather than assessed values), and the *P*'s and *Q*'s are the market rental value of per-student residential and non-residential property tax base in district *d* respectively. I calculate the *P*'s and *Q*'s by applying the $\tilde{\phi}$ and $\tilde{\gamma}$ factors to the average house value determined endogenously in the equilibrium computation.

In quantifying property tax bases for 2000 I face the challenge that since DPE no longer applied in 2000, I cannot use a similar approach to recover the adjustment factors. Having experimented with alternative mechanisms to obtain the 2000 adjustment factors and noticing the similarity of the results, I opted for the simplest solution of using the same as in 1990. In particular, I apply the $\tilde{\gamma}$ factors to the average house value determined endogenously in the equilibrium computation to calculate the *P*'s and *Q*'s.

[43] I apply the same user cost rate to 1990 and 2000 in order to focus on rental value changes due exclusively to house value changes. See Epple and Ferreyra (2007) for details on the calculation of the user cost rate.

Finally, it is necessary to compute the total metropolitan area income in order to calculate the equilibrium income tax rate. I compute total income as $Y = \sum_d \bar{y}_d \gamma_d n_d$, where the number of households in each district is adjusted by the spending factor γ_d for consistency with (A4). Without this adjustment, average tax burden per household would dramatically increase under Proposal A, which relies less on property taxes and more on income taxes. This would not be correct given that the property tax reform substituted tax instruments without a dramatic increase in tax burdens.