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

Maria Marta Ferreyra
Carnegie Mellon University, mferrey@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/tepper
Part of the Economic Policy Commons, and the Industrial Organization Commons

Published In
Although typically aiming at specific effects, large-scale policies can trigger other, general equilibrium effects. 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 policies but also for 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. Then, I apply the framework to study the effects of school finance reform on the Detroit metropolitan area.1 Concern about the equity

* Ferreyra: Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890 (e-mail: mferrey@andrew.cmu.edu). 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 Olszt and Glenda Rader from the government of the state of Michigan answered many of my questions on Proposal A and Michigan’s reforms in the 1990s, and facilitated my access to data. I am indebted to E. Anthon Eff, Dennis Epple, David Frame, Laurent Gobillon, Andrew Haughwout, Maurizio Mazzocco, Holger Sieg, Fallaw Sowell, and Steve Stern for valuable conversations and comments, and to participants at the 2008 Society for Economic Dynamics meetings, 2007 Regional Science Association meetings, 2004 Southern Economic Association meetings, and the 2004 summer workshop at the Federal Reserve Bank 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.

† To comment on this article in the online discussion forum visit the articles page at:

1 Throughout, I consider the metropolitan area as composed of Macomb, Oakland, and Wayne counties.
and adequacy of public school funding systems has led most states to overhaul them in the last 30 years. Thus, in 1994 the state of Michigan implemented a largely unexpected reform, known as Proposal A, which shifted public school funding away from local school districts and onto the state. School districts no longer determine their property tax rates and revenues. Instead, they receive a per-student allowance (“foundation allowance”) from the state. This scheme increased revenue for low-revenue districts and capped revenue for high-revenue districts. In addition Proposal A, implemented a tax reform by reducing property tax rates on owner-occupied housing and raising the state sales tax.

In metropolitan areas where households choose locations and schools jointly, a reform such as Proposal A may alter school revenues and property taxes, as well as households’ choices, housing prices, and the qualities of public schools. For instance, a district that benefits from lower property taxes and higher revenues may experience an increase in property values and an improvement in local public school quality. The district may also attract residents from other jurisdictions, which might improve the composition of the district’s student body, and thus its peer quality, and further raise the district’s 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 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 group uses calibrated models to investigate the equilibrium effects of school finance reform. 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 in the following ways. First, my computational representation of Detroit includes all actual 83 school districts and is therefore richer than previous studies. Second, I model revenues by using the actual state aid formulas rather than the simplifications invoked in previous estimation, which is critical in light of my interest in school finance reform. Third, while recent studies estimate Tiebout models with peer effects, I am the first to use school achievement data to identify the importance of peer quality (proxied by parental income) relative to school spending. 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. Thus, in this paper, I employ elements from the methodological frontier, enhance them, and apply them to an important 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. While useful, this type of exercise is uncommon, perhaps because regime shifts are rare (Michael Keane and Kenneth Wolpin 2006). My framework for large-scale policy analysis is similar to Holger Sieg et al. (2004) in the use and estimation of a Tiebout model. It differs, however, in that I allow the equilibrium level of the public good to adjust endogenously in the simulations, whereas they hold it fixed. Thus, my approach is particularly adept to study large-scale policies.

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, although some have studied general equilibrium effects from a reduced-form perspective. In particular, Dennis 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 (Rob Reich 2007). Hence, an important contribution in this paper is the equilibrium analysis of an adequate funding regime, an exercise not conducted in previous 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 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.

My analysis indicates that while Proposal A equalizes revenues to some extent, it is less effective at closing the school quality gap. This is 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 school quality, change little 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: what is the funding level necessary to secure the desired achievement in each district? I conduct an adequacy simulation in which the 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 solely on revenues. While other studies have reached a similar conclusion, this is the only paper to do so 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.

The remainder of this paper is organized as follows. Section I highlights some changes in the Detroit metropolitan area between 1990 and 2000. Section II presents the theoretical model, and Section III describes the model’s computational version. Section IV presents the estimation strategy and results. Section V presents the out-of-sample prediction exercise. Section VI presents the policy analysis, and Section VII concludes.

I. 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, public school spending per student, and public school quality. The exogenous variables are the state public school finance regime, the metropolitan area income distribution, the district-level stock of nonresidential 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

---

9 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.

10 See the discussion and references in Section VI.
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, with 83 school
districts and a population of about 3.93 million in 1990. Approximately one million
people 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 per-
tain to households with children in K–12 schools and come from the 1990 and 2000
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.11
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 rev-
enues 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 III 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 dis-
tricts by revenue, and that revenue changes were relatively small in the metropolitan
area, 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), for which the average pass
rate rose from 34 percent in 1991 to 70 percent in 2000.13

The metropolitan area income distribution also changed over the decade. While
all segments of the income distribution experienced real gains, these were great-
est for the high and low ends. For instance, at the deciles of the household income
distribution on which I focus for computational purposes—tenth, thirtieth, fiftieth,
seventieth, and ninetieth percentiles—real income grew by approximately 24, 9, 8,
12, and 12 percent, respectively.

11 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 term of the corresponding school year. Revenue, spending, and aid are per
pupil measures. The terms “revenue” and “spending” are used interchangeably.
12 Revenue changes were more pronounced for rural districts located outside the Detroit metropolitan area.
13 The size of these gains must be interpreted with caution, as achievement gains are often quite large when a
new test is introduced (Daniel Koretz 2002). The expansion of public school choice and the public school account-
ability implemented in Michigan over the 1990s might explain some of the achievement growth, although these
programs seem to have gained strength only after 2000 (Cullen and Loeb 2004; and Paul Courant, Cullen, and
Loeb 2003) and some still remain quite limited, as documented below. Furthermore, Roy (2003) provides evi-
dence that Michigan’s academic gains are much smaller when measured by federal tests rather than state tests.
Figure 1. Detroit Metropolitan Area in 1990

Figure 3 depicts 1990 neighborhood average housing quality. This varied considerably across neighborhoods, and neighborhoods in the central city had particularly low housing quality. Housing stock in the metropolitan area grew by 6.7 percent and most of the growth took place in the 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), improving most in the outer suburbs, although some neighborhoods in the central city improved as well.

To summarize, during the 1990s, 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.

II. 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

---

14 See Section III for the definition of neighborhood and the calculation of housing quality parameters.
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 the metropolitan area. Although 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 must 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:

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

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 \( \varepsilon \) is the household’s idiosyncratic preference for the location. For a given household, \( \varepsilon \) varies across locations. Furthermore, \( \varepsilon \) 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:

\[
c + (1 + t_d) p_{dh} = (1 - t_y) y_n + p_n,
\]

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:

\[
s = q^\rho x^{1-\rho},
\]

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, for example, parental engagement in the student’s and the school’s activities (Robert McMillan 2000), and parenting skills and home inputs in the production of achievement (Michael Rebell and Jessica 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:

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

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 nonresidential property in the district, respectively. The public school finance regime described in this section applies before the reform. In Section III, I note the modifications that apply after the reform.

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 a 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. Household preferences over property tax rates are single peaked, and property tax rates in each district are determined by majority voting.

The state cooperates with district $d$ by providing the per-student aid $AID_d$, funded by a state income tax of which rate $t_y$ balances the state budget constraint. The state applies a District Power Equalization (DPE) regime which guarantees a dollar yield per mill $^{17}$ 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:

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

which voters internalize when voting for local property taxes.

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 every house is occupied, property tax rates $t_d$ are consistent with majority voting by residents who take their location, property values, and the choices of others. Household preferences over property tax rates are single peaked, and property tax rates in each district are determined by majority voting.

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

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). The presence of private alternatives to public services violates single-peakedness of preferences over tax rates (Joseph 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 Nechyba (1999) and Calabrese et al. (2006) for further discussion and references.

Property tax rates are often expressed in mills. A one-mill rate is equivalent to a rate of 0.1 percent.
choices of others as given when voting on local tax rates; the budget balances for each district; the state budget balances; and at prices $p_{db}$, 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 based on equilibrium sufficient conditions.18

III. 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. Below, I outline this representation and the algorithm.

A. Community Structure and Households

My computational representation of the Detroit metropolitan area includes the actual number of districts. I construct neighborhoods so that the central city has the ten neighborhoods identified by the city’s actual classification of census tracts, 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 census tract-level data from the Detroit metropolitan area as follows. First, I associate census tracts and school districts using Geographic Information Systems (GIS) software. Then, I regress the logarithm of tract average rental price on tract-level neighborhood characteristics and school district fixed effects.19 I compute each tract’s quality index as the tract’s fitted rental value net of the school district fixed effect. This quality measure captures housing and amenities excluding school quality. Finally, I set each neighborhood’s quality index equal to the median index among the tracts 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 neighborhood differ because of the observed differences in

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.

19 The tract-level characteristics I include in the regression are average housing features from the census, and linear and quadratic terms in tract distance to the metropolitan area center.
housing and neighborhood characteristics. Since 36 percent of 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 tenth, thirtieth, fiftieth, seventieth, and ninetieth percentiles of the income distribution for households with children in grades K–12 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 that of the observed metropolitan area. Since households are assumed to differ in idiosyncratic location preferences, I assume that $\varepsilon$ 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 $\varepsilon$ equals $(1/6) \pi^2 b^2$.

B. 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. Furthermore, since DPE taxes residential and nonresidential property equally and Proposal A does not, I need to quantify each type of property separately to compare policy outcomes. The Web Appendix, available at http://www.aeaweb.org/articles.php?doi=10.1257/pol.1.1.147, describes the implicit formula and the property tax base quantification. Furthermore, under DPE, the state funded state aid largely through income and sales taxes. Since consumption, taxed by the sales tax, can be broadly viewed as proportional to income, I simplify throughout by considering only income taxes. 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 dollar amounts of foundation allowances ($fa$) for 1999 were determined as follows:

---

20 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 1990–1991 (C. Philip 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.

21 Epple and Ferreyra (2008) consider the separate role of sales and income taxes. Their main results are unchanged when only income taxes are considered.
(6)  
\[ fa = \begin{cases} 
5,700 & \text{if } x \leq 5,502 \\
0.833x + 1,114.34 & \text{if } x \in (5,502, 7,494) \\
0.867x + 860 & \text{if } x \geq 7,494.
\end{cases} \]

Under Proposal A, the state requires each district to levy 18 mills on nonresidential property and covers the difference between the foundation allowance and this local revenue through a 6-mill tax on residential and nonresidential 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.

If I had the exogenous data needed to compute the 1993 equilibrium, I would be able to predict the 2000 foundation allowances because these are a function of 1993 revenues, as indicated above. Since such data are not available, to compute the 2000 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. Moreover, a district with DPE spending below the maximum state guarantee has a Proposal A spending equal to its foundation allowance \( f_a_d \), and a state aid equal to \( AID_d = f_a_d - \bar{t}_Q Q_d/n_d \), where \( \bar{t}_Q \) is the required 18 mills on nonresidential property tax rate. Households in this district do not vote for property taxes. In contrast, a hold harmless district has spending \( x_d = \min (f_a_d, \bar{f} + t_d P_d/n_d) \), 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. This district receives state aid equal to \( AID_d = \bar{f} - \bar{t}_Q Q_d/n_d \).

Under Proposal A, the state’s budget constraint is  
\[ t_y y + \bar{t} \sum_d (P_d + Q_d) = \sum_d AID_d n_d, \]
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, for which 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

C. The Algorithm

In the model, the parameter vector is \( \theta = (\alpha, \beta, \rho, b) \). To compute the equilibrium for a given parameter point, the algorithm iterates as households choose locations 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, nonresidential property, metropolitan area income distribution, and state aid rule) and the initial distribution

---

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.
of household types and housing prices. The output is the computed equilibrium, which yields the predicted values of the variables of interest.\textsuperscript{23}

IV. Estimation

In the estimation, I match prereform 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 \( y_4 \) = fraction of public school students who pass the fourth-grade math test, normalized by the metropolitan area’s highest fraction (often called “school quality” in what follows).\textsuperscript{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 nonresidential property, quantity and quality of housing, and data pertaining to the metropolitan area (tenth, thirtieth, fiftieth, seventieth and ninetieth income percentiles, and school funding regime). I assume the following:

\begin{equation}
E(y_{ji} \mid X_i) = h_j(X_i, \theta) \quad j = 1, \ldots, 4; i = 1, \ldots, D,
\end{equation}

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 \( \theta \). 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 \( V(y_{ji} \mid X_i) = \sigma_{jj}/n_i = \sigma_{jji}/n_i = \sigma_{jji}^2 = \sigma_{jji} \), where \( \sigma_{jk} \) and \( \sigma_{jji}^2 \) denote population covariances and variances, respectively.

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

\begin{equation}
L(\theta) = \sum_{j=1}^{4} \sum_{i=1}^{D} (y_{ij} - \hat{y}_{ij}(\theta))^2.
\end{equation}

I use the residuals from this stage to compute the \( \hat{\sigma}_{jki} \)'s needed to account for heteroskedasticity and cross-equation covariances. In the second stage, I minimize the following loss function:

\begin{equation}
\tilde{L}(\theta) = \sum_{j=1}^{4} \sum_{k=1}^{4} \sum_{i=1}^{D} \hat{\sigma}_{jki} (y_{ji} - \hat{y}_{ji}(\theta)) (y_{ki} - \hat{y}_{ki}(\theta)),
\end{equation}

\textsuperscript{23} More details on the algorithm can be found in the appendix to Ferreyra (2007).

\textsuperscript{24} The school quality predictions are also normalized by the highest predicted quality in the metropolitan area in order to fit pass rates, which are 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.
where each $\hat{\sigma}_{jki}$ is a function of the corresponding $\hat{\sigma}_{jki}$'s. The value of $\theta$ 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. Spending and housing prices identify the school quality coefficient on the utility function ($\alpha$), as a higher $\alpha$ raises spending and most housing prices. Housing prices also identify the consumption coefficient in the utility function ($\beta$) because a higher $\beta$ 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 ($\rho$), as a higher $\rho$ 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. A greater $b$ makes household sorting depend less on income and more on idiosyncratic preferences, and 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 $\rho$ implies that peer quality contributes, on the margin, more 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–D 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 (0.84, 0.88, 0.76, and 0.83 for income, rental value, spending and school quality, respectively) indicate a reasonably good fit for 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 underpredicted because the empirical income distribution is truncated at the ninetieth 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 $\rho$.

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 (Thomas...

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.
Overall, I view the evidence presented here as indicative that the model successfully captures patterns in the data.

V. 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 2000 data. Some issues arise when predicting the 2000 equilibrium. For
instance, pass rates rose consistently during the 1990s, perhaps due to learning about the test (see Section I). This improvement displayed a ceiling effect; low-performing districts, with the greatest room for gains, 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 consistent with the ceiling effect.26 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.

Table 1—Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.137</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>β</td>
<td>0.740</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>b</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Sum of squared residuals</td>
<td>569.669</td>
</tr>
</tbody>
</table>


Table 2—Root Mean Squared Error for In- and Out-of-Sample Predictions

<table>
<thead>
<tr>
<th></th>
<th>In-sample</th>
<th>Out-of-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>1.89</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>(6.44)</td>
<td>(7.39)</td>
</tr>
<tr>
<td>Rental value</td>
<td>0.40</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>(1.26)</td>
<td>(2.02)</td>
</tr>
<tr>
<td>Spending</td>
<td>0.17</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.64)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>School quality</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(0.77)</td>
</tr>
</tbody>
</table>

Notes: Observations: 83 school districts. Data refer to district averages and are weighted by district number of housing units. Observed sample means are in parentheses. Income, rental value, and spending are expressed in $10,000. School quality ranges between zero and one. Data are weighted by district number of housing units.

26 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. 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 the next section.
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 underpredicts the corresponding DPE revenues. In contrast, the model overpredicts 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. 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 0.71). Figure 6E displays observed and fitted school quality. The model predicts school quality quite well (the correlation between observed and predicted values is 0.69), although with underprediction for some medium-performance districts and slight overprediction for top districts.

A comparison of panel A in Tables 3 and 4 reveals that correlations among the variables of interest changed little 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 replicates the observed 2000 correlations reasonably well.

One might ask whether the model captures changes in the endogenous variables during the decade. 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 toward lower-income locations in response to Proposal A’s

---

**Table 3—In-Sample Goodness of Fit: Some Correlations**

<table>
<thead>
<tr>
<th></th>
<th>Income</th>
<th>Rental value</th>
<th>Spending</th>
<th>School quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Observed data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rental value</td>
<td>0.98</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spending</td>
<td>0.65</td>
<td>0.63</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>School quality</td>
<td>0.85</td>
<td>0.86</td>
<td>0.59</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Panel B: Fitted data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rental value</td>
<td>0.99</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spending</td>
<td>0.46</td>
<td>0.44</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>School quality</td>
<td>0.99</td>
<td>0.98</td>
<td>0.54</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Notes: Observations: 83 school districts. Data refer to district averages, and correlations are weighted by district number of housing units.
Figure 6. Out-of-Sample Goodness of Fit: Fitted versus Observed Values

Notes: See Figure 5. Correlations between fitted and observed values are weighted by the observation’s measure of households and are as follows: 0.82, 0.86, 0.71, 0.71, and 0.69 for Figures 6A–E, respectively.
incentives. Alternatively, since the low segment of the income distribution experienced the greatest proportional gains in real income (see Section I), low-income districts might have grown richer simply because their original inhabitants became richer. I reexamine this matter in Section VI.

Moreover, rental values grew 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

<table>
<thead>
<tr>
<th>Income</th>
<th>Rental value</th>
<th>Spending</th>
<th>School quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Observed data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rental value</td>
<td>0.99</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Spending</td>
<td>0.50</td>
<td>0.48</td>
<td>1.00</td>
</tr>
<tr>
<td>School quality</td>
<td>0.69</td>
<td>0.69</td>
<td>0.46</td>
</tr>
</tbody>
</table>

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

Table 5—Out-of-Sample Goodness of Fit: Correlation between Observed and Predicted Changes

<table>
<thead>
<tr>
<th>Income</th>
<th>Rental value</th>
<th>Revenue</th>
<th>School quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.39</td>
<td>0.46</td>
<td>0.60</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Notes: Observations: 83 districts. Data refer to the correlation between district predicted and observed change in the corresponding average. Data are weighted by district number of housing units.

Table 6—Out-of-Sample Goodness of Fit: Correlations between 1990 Values and Percent Changes

<table>
<thead>
<tr>
<th>Income</th>
<th>Rental value</th>
<th>Revenue</th>
<th>School quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>−0.23</td>
<td>−0.63</td>
<td>−0.72</td>
</tr>
<tr>
<td>Rental value</td>
<td>−0.26</td>
<td>−0.72</td>
<td>−0.76</td>
</tr>
</tbody>
</table>

Notes: 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. Data are weighted by district number of housing units.
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 negligible outside the central city (and quite small within). 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 summarize, the model fits the out-of-sample data reasonably well. This lends plausibility to the counterfactual 2000 DPE and provides confidence for policy analysis.\textsuperscript{28}

VI. 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 V.

\begin{table}[h]
\centering
\caption{Out-of-Sample Goodness of Fit: Correlations between Changes}
\begin{tabular}{lcccccc}
\hline
 & Income & Rental value & Spending & School quality & Housing quality & Relative size \\
\hline
Panel A: Observed data & & & & & & \\
Income & 1.00 & & & & & \\
Rental value & 0.49 & 1.00 & & & & \\
Spending & 0.13 & 0.27 & 1.00 & & & \\
School quality & 0.22 & 0.59 & 0.20 & 1.00 & & \\
Housing quality & 0.35 & 0.20 & 0.26 & 0.05 & 1.00 & \\
Relative size & \(-0.35\) & \(-0.72\) & \(-0.03\) & \(-0.73\) & 0.05 & 1.00 \\
Panel B: Fitted data & & & & & & \\
Income & 1.00 & & & & & \\
Rental value & 0.75 & 1.00 & & & & \\
Spending & 0.47 & 0.40 & 1.00 & & & \\
School quality & 0.67 & 0.96 & 0.50 & 1.00 & & \\
Housing quality & 0.48 & 0.25 & 0.27 & 0.09 & 1.00 & \\
Relative size & \(-0.17\) & \(-0.67\) & 0.21 & \(-0.56\) & 0.05 & 1.00 \\
\hline
\end{tabular}
\textit{Notes:} 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.
\end{table}

\textsuperscript{28} The out-of-sample data corresponds to a setting where school districts have little discretion to determine expenditure. This might appear as a limitation of 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, as is the case here.
Table 8 characterizes the equilibrium for the benchmark DPE and alternative policies by presenting effects on school revenues, demographics, property values, school 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 largely by households in high-income districts. Variation (measured by the ratio of the highest to the lowest value and the ratio of the seventy-fifth to the twenty-fifth 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 average net income tax subsidy per student (state aid per student minus the household’s income tax liability) is positive for them and negative for wealthier districts.

A. Proposal A

Column 2 of Table 8 displays the effects of Proposal A. Figure 7 depicts Proposal A’s revenues relative to the benchmark 2000 DPE. As is clear from (6), the 25 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 Table 8, 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 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 to those with relatively good housing. As these relocations take place, income variation drops across districts. 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 during the 1990s (see Section V) are more likely associated with changes in the overall income

29 In these simulations, “high-income districts” are those allowed to raise hold-harmless mills in Proposal A.
30 In the case of income, the highest to lowest ratio is the same for DPE and the other revenue-equity policies, whereas the seventy-fifth to twenty-fifth percentile ratio is lower for Proposal A and the other policies than it is 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.
distribution, which were relatively more favorable for original residents in those districts, than with household relocations.

Changes in property values are similar to changes in income. Because housing price changes reflect the net impact of the property tax and revenue reform, 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 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 10 percent revenue increase is associated, on average, with about two additional percentage points in the pass rate, similar to estimates reported by Papke (2005).

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 nonresidential taxes, which are a subsidy to households, and to higher income taxes. In contrast to the benchmark, students in high-income districts receive some state aid under Proposal A, although these households also pay larger income tax bills. Hence, their net income tax liability increases from $500 to $4,000. The reverse is true for students in other districts, whose average income tax subsidy rises from $100 to $1,100. Although high-income districts undergo, on average, roughly 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 (John 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.31

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.

---

31 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.
## Table 8—Effects of Alternative Policies

<table>
<thead>
<tr>
<th>Proposal</th>
<th>DPE</th>
<th>Tax reform</th>
<th>Low GTB</th>
<th>High GTB foundation</th>
<th>Low foundation</th>
<th>High foundation</th>
<th>Adequacy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Revenue per student</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average revenue</td>
<td>$5,700</td>
<td>$6,400</td>
<td>$5,700</td>
<td>$6,100</td>
<td>$9,900</td>
<td>$6,100</td>
<td>$14,900</td>
</tr>
<tr>
<td>Urban districts</td>
<td>$3,600</td>
<td>$5,700</td>
<td>$3,600</td>
<td>$4,500</td>
<td>$8,800</td>
<td>$6,100</td>
<td>$14,900</td>
</tr>
<tr>
<td>Suburban districts</td>
<td>$6,400</td>
<td>$6,700</td>
<td>$6,400</td>
<td>$6,600</td>
<td>$10,200</td>
<td>$6,100</td>
<td>$14,900</td>
</tr>
<tr>
<td>High-income districts</td>
<td>$10,100</td>
<td>$8,200</td>
<td>$10,100</td>
<td>$10,100</td>
<td>$13,000</td>
<td>$6,100</td>
<td>$14,900</td>
</tr>
<tr>
<td>Highest / lowest</td>
<td>10.06</td>
<td>2.30</td>
<td>10.06</td>
<td>7.51</td>
<td>4.22</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>75th percentile / 25th percentile</td>
<td>2.21</td>
<td>1.26</td>
<td>2.21</td>
<td>1.84</td>
<td>1.28</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Average proportional change</td>
<td>0.34</td>
<td>0.00</td>
<td>0.10</td>
<td>0.96</td>
<td>0.35</td>
<td>2.28</td>
<td>5.10</td>
</tr>
<tr>
<td><strong>Panel B: Household income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average revenue</td>
<td>$66,700</td>
<td>$66,700</td>
<td>$66,700</td>
<td>$66,700</td>
<td>$66,700</td>
<td>$66,700</td>
<td>$66,700</td>
</tr>
<tr>
<td>Urban districts</td>
<td>$26,700</td>
<td>$26,900</td>
<td>$26,700</td>
<td>$25,700</td>
<td>$21,100</td>
<td>$27,100</td>
<td>$27,000</td>
</tr>
<tr>
<td>Suburban districts</td>
<td>$79,100</td>
<td>$79,100</td>
<td>$79,100</td>
<td>$79,400</td>
<td>$79,000</td>
<td>$79,000</td>
<td>$79,000</td>
</tr>
<tr>
<td>High-income districts</td>
<td>$94,100</td>
<td>$91,000</td>
<td>$93,800</td>
<td>$91,000</td>
<td>$88,300</td>
<td>$89,300</td>
<td>$89,300</td>
</tr>
<tr>
<td>Highest / lowest</td>
<td>8.40</td>
<td>8.40</td>
<td>8.40</td>
<td>8.40</td>
<td>8.40</td>
<td>8.40</td>
<td>8.40</td>
</tr>
<tr>
<td>75th percentile / 25th percentile</td>
<td>2.54</td>
<td>2.42</td>
<td>2.53</td>
<td>2.50</td>
<td>2.49</td>
<td>2.34</td>
<td>2.34</td>
</tr>
<tr>
<td>Average proportional change</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Panel C: Rental value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average rental revenue</td>
<td>$21,500</td>
<td>$21,900</td>
<td>$22,300</td>
<td>$21,300</td>
<td>$20,900</td>
<td>$21,800</td>
<td>$21,100</td>
</tr>
<tr>
<td>Urban districts</td>
<td>$12,100</td>
<td>$12,400</td>
<td>$12,600</td>
<td>$12,100</td>
<td>$11,900</td>
<td>$12,400</td>
<td>$12,300</td>
</tr>
<tr>
<td>Suburban districts</td>
<td>$24,400</td>
<td>$24,800</td>
<td>$25,300</td>
<td>$24,200</td>
<td>$23,700</td>
<td>$24,700</td>
<td>$23,900</td>
</tr>
<tr>
<td>High-income districts</td>
<td>$28,200</td>
<td>$27,800</td>
<td>$28,700</td>
<td>$27,800</td>
<td>$26,500</td>
<td>$26,900</td>
<td>$26,000</td>
</tr>
<tr>
<td>Highest / lowest</td>
<td>4.49</td>
<td>4.48</td>
<td>4.57</td>
<td>4.46</td>
<td>4.32</td>
<td>4.45</td>
<td>4.27</td>
</tr>
<tr>
<td>75th percentile / 25th percentile</td>
<td>2.05</td>
<td>1.97</td>
<td>2.00</td>
<td>2.02</td>
<td>1.98</td>
<td>1.91</td>
<td>1.88</td>
</tr>
<tr>
<td>Average proportional change</td>
<td>0.02</td>
<td>0.04</td>
<td>−0.01</td>
<td>−0.02</td>
<td>0.02</td>
<td>−0.01</td>
<td>−0.08</td>
</tr>
<tr>
<td><strong>Panel D: School quality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average school quality</td>
<td>0.50</td>
<td>0.52</td>
<td>0.51</td>
<td>0.51</td>
<td>0.52</td>
<td>0.54</td>
<td>0.53</td>
</tr>
<tr>
<td>Urban districts</td>
<td>0.22</td>
<td>0.23</td>
<td>0.22</td>
<td>0.22</td>
<td>0.23</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Suburban districts</td>
<td>0.59</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.61</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>High-income districts</td>
<td>0.73</td>
<td>0.70</td>
<td>0.73</td>
<td>0.73</td>
<td>0.72</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Highest / lowest</td>
<td>8.21</td>
<td>6.81</td>
<td>8.20</td>
<td>7.89</td>
<td>7.40</td>
<td>6.38</td>
<td>6.38</td>
</tr>
<tr>
<td>75th percentile / 25th percentile</td>
<td>2.27</td>
<td>2.22</td>
<td>2.26</td>
<td>2.25</td>
<td>2.22</td>
<td>2.10</td>
<td>2.10</td>
</tr>
<tr>
<td>Average proportional change</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.11</td>
<td>0.11</td>
<td>0.22</td>
</tr>
<tr>
<td>Urban districts</td>
<td>0.08</td>
<td>0.01</td>
<td>0.03</td>
<td>0.05</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>Suburban districts</td>
<td>0.04</td>
<td>0.00</td>
<td>0.01</td>
<td>0.07</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>High-income districts</td>
<td>−0.04</td>
<td>0.00</td>
<td>−0.01</td>
<td>−0.03</td>
<td>−0.06</td>
<td>−0.06</td>
<td>−0.06</td>
</tr>
</tbody>
</table>
tax to fund K–12 schools, as was the case in California and New Hampshire (Therese McGuire and Leslie Papke 2007). 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.\(^{32}\) Hence, the reform leads to overall housing appreciation (4 percent on average), particularly in urban districts which have the highest benchmark property tax rates. Of all the policies

\(^{32}\) 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 nonresidential property taxes creates a subsidy for households, which 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.

Table 8—Effects of Alternative Policies (Continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential property tax rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.07</td>
<td>0.02</td>
<td>0.03</td>
<td>0.07</td>
<td>0.07</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.09</td>
<td>0.06</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Nonresidential property tax rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.07</td>
<td>0.10</td>
<td>0.10</td>
<td>0.07</td>
<td>0.07</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.04</td>
<td>0.10</td>
<td>0.10</td>
<td>0.04</td>
<td>0.04</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
<td>0.08</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Income tax rate</td>
<td>0.00</td>
<td>0.20</td>
<td>0.01</td>
<td>0.00</td>
<td>0.03</td>
<td>0.02</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>Average tax burden per household</td>
<td>$1,700</td>
<td>$1,800</td>
<td>$1,400</td>
<td>$1,800</td>
<td>$3,300</td>
<td>$1,600</td>
<td>$5,000</td>
<td>$6,800</td>
</tr>
<tr>
<td>Average net income tax subsidy per student</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
</tr>
<tr>
<td>High-income districts</td>
<td>−$500</td>
<td>−$4,000</td>
<td>−$2,600</td>
<td>−$1,100</td>
<td>−$4,500</td>
<td>−$5,100</td>
<td>−$11,100</td>
<td>−$25,700</td>
</tr>
<tr>
<td>Low-income districts</td>
<td>$100</td>
<td>$1,100</td>
<td>$700</td>
<td>$300</td>
<td>$1,200</td>
<td>$1,400</td>
<td>$3,100</td>
<td>$6,600</td>
</tr>
<tr>
<td>Average share of residential property taxes</td>
<td>0.69</td>
<td>0.21</td>
<td>0.30</td>
<td>0.64</td>
<td>0.38</td>
<td>0.22</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>Average share of non-residential property taxes</td>
<td>0.22</td>
<td>0.26</td>
<td>0.31</td>
<td>0.21</td>
<td>0.13</td>
<td>0.30</td>
<td>0.12</td>
<td>0.21</td>
</tr>
<tr>
<td>Average share of income taxes</td>
<td>0.09</td>
<td>0.53</td>
<td>0.39</td>
<td>0.15</td>
<td>0.49</td>
<td>0.48</td>
<td>0.79</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Notes: Observations: 83 districts. For a given variable, “Highest/Lowest” is the ratio of the metropolitan area’s highest to lowest district average; “75th percentile/25th percentile” 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 are rounded to closest hundred. All averages are weighted. Weight is number of housing units.
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. 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 (Yinger 2004).

To summarize, one goal of Proposal A was property tax reduction. Another goal 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 DPE and foundation formulas (Yinger 2004), which I analyze below. The benchmark for comparison continues to be the 2000 DPE.

B. District Power Equalization

As of 1993 Michigan already had a DPE regime aimed at equalizing revenues. One reason revenues varied so much across districts, despite the DPE regime, was that policymakers 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 achieve 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 thirtieth percentile of the metropolitan area distribution of property tax base per child to slightly above the fortieth 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 14 in-formula districts receive at least twice as much aid as in the benchmark, and 10 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 because it preserves local discretion. Since the low GTB policy alters revenues less than Proposal A, demographic and school quality changes are also smaller.

William Duncombe and Yinger (1998) discuss the limitations of DPE to achieve revenue equalization. Quantifying these 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 through a very high GTB. Thus, column 5 of Table 8 explores the effects of increasing the initial GTB by a factor of 5, which amounts to a real GTB increase of about 270 percent and leaves GTB at the ninety-ninth percentile of the distribution of property tax base per student. This regime (“high GTB”) amounts to an almost complete equalization of the property tax base per student, with 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, 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 ninety-fifth percentile of the distribution of property tax base per student), yet revenues still varied across districts.

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

---

**Figure 7. Log of Per-Student Revenue under Alternative Regimes**

*Notes:* “Log” stands for natural logarithm. Revenues are expressed in $10,000. The 45-degree line represents the simulations’ benchmark (2000 DPE).

*Source:* Author’s simulations.
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.

C. 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. 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 regime induces greater demographic changes than Proposal A, the changes are 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 does not fall as much as revenue variation, showing 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 element on which households sort. Since housing qualities are the same in both programs, so are households’ choices and school qualities.

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, heavy reliance on income taxes, and high tax burden. Despite its high fiscal cost, the high foundation remains unable to eliminate the achievement gap.

---

33 Yinger (2004) and William Fischel (2001) review the literature that explores this issue. Fischel (2001) examines individual states and concludes that leveling down may have prevailed.

34 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.
Foundations’ inability to equalize achievement is consistent with the 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.

D. Adequacy

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). Plaintiffs have adopted a variety of approaches to determine the cost of the desired achievement (Thomas Downes and Leanna Stiefel 2007). 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 might attract higher-income households 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 for which the 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, 10 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 funding required for the adequacy program amounts to an income tax rate close to 65 percent. However, this dramatic revenue increase for low-performing districts generates some household relocation towards these districts, which in turn improve their peer qualities. Hence, 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 very fiscally 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, largely 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 10 other districts. Since state aid is funded largely through 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 seventy-fifth to twenty-fifth school quality percentiles does not change much, the 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 2013–2014, my simulations suggest that the cost of attaining this goal merely through revenue policies would be nothing short of prohibitive.35

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, nonrevenue based policies might succeed. Public school accountability is one such policy which seems to have been particularly effective among low-performing districts and students (Brian Jacob 2005, Hanley Chiang 2007, and Cecilia Rouse et al. 2007). As a result, a number of schools have implemented positive behavioral changes (Chiang 2007, Rouse et al. 2007, and Rebell and Wolff 2008). These effects are noteworthy, particularly because accountability’s cost is negligible relative to that of school funding reform (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.36 Perhaps more importantly, my measure of peer quality (household income) is correlated with the quality of parental

35 Reich (2007, page 18) 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. Jennifer Imazeki (2007) and Jon Sonstelie (2007) 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, Goertz, Odden 2007).

36 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 participated 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 details on the public school choice programs in Michigan.
and home inputs in the production of child achievement. The recognition that socioeconomic resources are related to achievement hardly provides a policy prescription to close the achievement gap (Greg Duncan and Katherine Magnuson 2005). However, some high-quality interventions engaging children and their parents, particularly before school entry, have proved to be successful (Loeb and Daphna Bassok 2007). Furthermore, recent research highlights that the cognitive skills targeted in school reform programs are complementary with noncognitive skills (such as character traits and personal habits) in the production of human capital (Flavio 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.

VII. 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 a 1994 reform (Proposal A). 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 the model for policy analysis. According to my simulation, closing the achievement gap across districts only by means of extra funding is prohibitively costly, even after taking into account mobility effects.

My study relies on a short-run model that assumes 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 strongly related to other changes, as noted in the discussion of the out-of-sample exercise. 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. 2004, Calabrese et al. 2006, Ferreyra 2007, and the references therein), and future research that might endogenize the housing stock.

37 The importance of home inputs is highlighted by Todd and Wolpin (2007), who estimate that approximately 50 percent of the gap is attributable to differences in mothers’ pre-market skills, and that between 10 and 20 percent of the black-white achievement gap could be closed by equalizing home inputs.
38 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 that occurred in the United States as a whole during the 1990s (David Autor, Lawrence Katz, and Melissa Kearney 2005).
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.

REFERENCES


