Chapter 5.2: Property Tax Capitalization

Introduction

A famous article by Oates (1969) first tested the hypothesis that house values depend on local public service quality and on local property tax rates, a phenomenon known as capitalization.\(^1\) His strong evidence for capitalization in a sample of communities in New Jersey stimulated dozens of additional studies on the topic.

This chapter covers the subsequent literature on property tax capitalization. It focuses on the approach and evidence in Yinger, et al. (1988), henceforth referred to as \textit{PTHV}. The chapter considers conceptual and empirical issues that arise in estimating property tax capitalization and discusses the important, but often counter-intuitive implications of property tax capitalization for public policy.

The Theory of Property Tax Capitalization

Property tax capitalization arises from the basic equality between the value of an asset, in this case a house, and the present value of the net benefits from owning it. Without property taxes, the amount someone is willing to pay for a house is the present value of the rental benefits from owning it, or

\[
V = \sum_{y=1}^{L} \frac{\hat{P}_H}{(1+r)^y} = \frac{\hat{P}_H}{(1+r)} + \frac{\hat{P}_H}{(1+r)^2} + \ldots + \frac{\hat{P}_H}{(1+r)^L} \tag{1}
\]

\(^1\) Oates was not the first to look at property tax capitalization, but he was the first to estimate tax and service capitalization in one unified framework. Earlier citations on property tax capitalization are provided by Yinger, et al. (1988).
where \( \hat{P} \) is the price of housing services not considering taxes, \( H \) is housing services (quality adjusted square feet), \( r \) is the real discount rate, and \( L \) is the expected lifetime of a house.

This equation can be simplified considerably with some simple algebra:

\[
V(1+r) = \hat{P}H + \frac{\hat{P}H}{(1+r)} + \frac{\hat{P}H}{(1+r)^2} + \ldots + \frac{\hat{P}H}{(1+r)^{L-1}}
\]

and

\[
V - V(1+r)^{-1} = \hat{P}H - \frac{\hat{P}H}{(1+r)^{L}}
\]

or

\[
V[1-(1+r)^{-1}] = \hat{P}H[(1+r)^{-L} - 1]
\]

or

\[
V = \hat{P}H \left( \frac{1-(1+r)^{-L}}{r} \right) = \frac{\hat{P}H}{r'}
\]

where

\[
r' = \frac{r}{1-(1+r)^{-L}}
\]

This derivation leads to a very important simplification. If the value of rental services is constant (in real terms) over time and the lifetime of housing, \( L \), is large, then

\[
r = r'
\]

and

\[
V = \frac{\hat{P}H}{r}
\]
Now we can add property taxes.

\[ V = \sum_{y=1}^{L} \frac{\hat{PH}}{(1+r)^y} - \sum_{y=1}^{N} \frac{\beta' tV}{(1+r)^y} \]

In this equation, \( \beta' \), is the "degree of property tax capitalization," defined as the impact of a $1 increase in the present value of the property tax stream on the value of a house. If \( \beta' = 0.5 \) for example, then a $1 increase in the present value of property taxes leads to a $0.50 decrease in the value of a house.

The reason for the prime will be clear in a moment.

Now you can see the source of the term "capitalization." The annual flow of property tax payments shows up in a "capital" or asset value, namely \( V \), using the logic of discounting. Indeed, the denominator in an asset pricing expression, \( r \) or \( r' \) here, is often called the capitalization rate.

Note also that we do not require the expected lifetime of property taxes to be the same as the expected lifetime of the house. The reason for this will become clear shortly.

As you can see, equation (4) is not in final form because \( V \) is on both sides. So let's solve for \( V \).

\[ V \left(1 + \frac{\beta' t}{r'} \right) = \frac{PH}{r} \]

or

\[ V = \frac{\hat{PH}}{r \left(1 + \frac{\beta' t}{r'} \right)} = \frac{\hat{PH}}{r + \left(\frac{r'}{r'} \beta' \right) t} = \frac{\hat{PH}}{r + \beta' [1 - (1+r)^{-N}] t} \]

\[ = \frac{\hat{PH}}{r + \beta t} = \frac{\hat{PH}}{r \left(1 + \frac{\beta}{r} t \right)} \]
The coefficient to be estimated, $\beta$, is the expression in front of $t$, so it includes both $\beta'$ and the impact of different expectations about the lifetime of a house and of property taxes. This explains why we need both $\beta$ and $\beta'$; the first is what we estimate but the second is the underlying degree of capitalization. It is still not obvious why we need to consider expectations—hold on.

**Estimating Property Tax Capitalization**

Although equation (5) is fairly simple, it has proven to be difficult to estimate for several reasons.

First, it involves a non-linear relationship between $t$ and $V$, even after taking logarithms, so it cannot be estimated with linear regression methods. As a result, existing studies use various approximations or non-linear estimating techniques.

If we take logs, the estimating equation is

$$
\ln(V) = \ln(\hat{P}) + \ln(H) - \ln(1 + \frac{\beta}{r} t)
$$

(6)

In this equation, the first term is a function of locational characteristics of housing and the second term is a function of the structural characteristics of housing. Because $r$ does not vary across houses, it is just the constant term.

More importantly, it is impossible to separate $\beta$ and $r$. I’ll return to this point below. Moreover, the ratio $\beta/r$, which is the coefficient of interest, enters in a nonlinear way. One possible approach is to estimate equation (6)—and $\beta/r$—using nonlinear methods.

Another possibility is to use the standard approximation that $\ln(1+a) \approx a$ when $a$ is small, say 0.1. With this approximation, the appropriate tax variable is simply $t$ by itself (and its estimated coefficient is $\beta/r$). The accuracy of this approximation depends on the value of $a = \beta t/r$. Suppose $t = 0.01$, $\beta = 0.3$, and $r = .03$. In this case, $a = 0.1$, and this approximation is reasonably accurate. But if $t$ or $\beta$ is any larger, then the quality of this approximation may not be very good. Moreover, a researcher cannot know the value of $\beta$ ahead of time, and therefore cannot know if this is a reasonable approach to take.
Nevertheless, many studies simply enter $t$ or $\ln(t)$ as the tax variable. The use of $\ln(t)$, the approach in Oates, is particularly troubling because it has no connection to equation (6). The errors introduced by these inaccurate approximations are not known. In my view, this approach is very unsatisfactory: The theory here is very clear and empirical work should make use of it.

Another approach, which is possible with panel data containing at least two time periods, is to use a difference form of the equation. With two time periods,

$$V_1 = \frac{\hat{PH}}{r + \beta t_1} \quad \text{and} \quad V_2 = \frac{\hat{PH}}{r + \beta t_2}$$

so

$$V_1 - V_2 \equiv \Delta V = \frac{\hat{PH}}{r + \beta t_1} - \frac{\hat{PH}}{r + \beta t_2} = \hat{PH} \left( \frac{1}{r + \beta t_1} - \frac{1}{r + \beta t_2} \right)$$

$$= \hat{PH} \left( \frac{\beta(t_2 - t_1)}{(r + \beta t_1)(r + \beta t_2)} \right) = \hat{PH} \left( \frac{\beta \Delta t}{r + \beta t_1} \right) \left( \frac{1}{r + \beta t_2} \right) = V_1 \left( \frac{\beta \Delta t}{r + \beta t_2} \right)$$

so

$$\frac{\Delta V}{V_1} = \left( \frac{\beta \Delta t}{r + \beta t_2} \right)$$

(7)

This equation can be estimated with nonlinear methods, where, as before, the coefficient is $\beta/r$.

In PTHV, this equation is used to study intra-jurisdictional capitalization, that is, the capitalization of effective property tax rate differences within a community. To remove the impact of intra-jurisdictional tax differences/changes, and of other factors that vary over time, the dependent variable is deflated using a housing price index. This removes from $V$ the impact of any change in the average effective tax rate (among other things) and leaves just the impact of the change in the deviation from the average tax rate.
To be more specific, let us express the effective tax rate as \( t = \bar{t} + (t - \bar{t}) \) and put an asterisk on \( V \) to remind us that it has been deflated using a price index. Then equation (7) becomes

\[
\frac{\Delta V^*}{V_1^*} = \frac{-\beta \Delta t}{r + \beta t_2} = \frac{-\beta \left( (t_2 - \bar{t}) + [(t_2 - \bar{t}) - (t_1 - \bar{t})] \right)}{r + \beta [\bar{t}_2 + (t_2 - \bar{t}) + (t_2 - \bar{t})]}
\]

Equation (8)

The price-index adjustment implies that \((\bar{t}_2 - \bar{t}_1) = 0\); in other words, the impact of district-level changes in effective tax rates has been removed from the dependent variable. In addition, revaluation, if it is done carefully, leads to an outcome in which every house has approximately the same effective tax rate after revaluation, which implies that \((t_2 - \bar{t}_2) = 0\). Introducing these two equations into equation (8) results in the following simple form:

\[
\frac{\Delta V^*}{V_1^*} = -\beta \left( (t_2 - \bar{t}) - (t_1 - \bar{t}) \right) = -\beta \Delta t^* = b \Delta t^*
\]

Equation (9)

where

\[
b = \frac{-\beta}{r + \beta \bar{t}_1}
\]

Equation (10)

Equation (10) provides an interpretation for the estimated coefficient, \( b \). Solving equation (10) for \( \beta \), we find that

\[
\beta = \frac{br}{1 - b \bar{t}_1}
\]

Equation (11)

Thus, with an estimate of \( b \), data on \( \bar{t}_1 \), and an assumption about the \( r \), one can obtain an estimate of \( \beta \).

\[\text{Note:} \quad \text{Note that it is possible to drop the } (t_2 - \bar{t}_2) \text{ term in the numerator of (9), but this approximation is not necessary since the change in } t \text{ can be observed. In addition, it is possible to use nonlinear methods to estimate (8) with this term in the denominator, but keeping the assumption that } (\bar{t}_2 - \bar{t}_1). \quad \text{See PTHV. These nonlinear results are virtually identical to the results obtained with equation (9).} \]


Equation (11) corrects an error in PTHV. The algebra in that book mistakenly ignores the $bT_i$ term in the denominator of (11). As we will see below, this mistake implies that the book understates the value of $\beta$ by about 30 percent.

Second, the value of the discount rate, $r$, is not observed, and the form of equations (6), (7), and (8) precludes separate estimation of $r$ and $\beta$. Most studies follow Oates by estimating a value of $\beta/r$, assuming a value for $r$, and then calculating the implied value of $\beta$. The trouble with this approach is that the value of $\beta$ depends on an untested assumption and different studies use different values of $r$. In fact, the most extreme estimates in the literature, in either direction, are driven largely by extreme assumptions about $r$.

Moreover, scholars are amazingly careless about $r$, often using a nominal interest rate, when the theory clearly shows that a real rate, say 3 to 5 percent, is needed.

A real rate equals the nominal or market rate minus anticipated inflation. PTHV starts with a long-run, low-risk nominal rate (as for an investment in housing) and then subtracts anticipated inflation based on a study of the factors that determine inflation expectations. This leads to a 3 percent real discount rate.

Third, the asset-pricing logic behind equations (6), (7), and (9) requires assumptions about house buyers’ expectations. To be specific, this logic predicts that a $1 increase in the present value of future property taxes will lead to a $1 decline in house value (i.e. $\beta' = 1$), but it does not say that current tax differences will be fully capitalized (i.e. $\beta = 1$) if they are not expected to persist.

Virtually all the literature estimates the capitalization of current property tax differences. Under the assumption that current tax differences will persist indefinitely, the assumption that $\beta = 1$ makes perfect sense. In fact, however, current differences may not be expected to persist. In this case, we can use the result derived earlier, namely,

$$\beta = \beta'[1-(1+r)^{-N}]$$  \hspace{1cm} (12)

where $N$ is the length of time current tax differences are expected to persist. The theory indicates that $\beta' = 1$, but the estimated $\beta$ clearly need not equal 1, and indeed need not equal the same value under all circumstances.
For example, if current property tax differences across (or within) communities are expected to disappear in 10 years and $r = .03$, then (3) implies that the estimated $\beta$ will be only 26 percent even if $\beta' = 1$.

Fourth, because $t=T/V$, one must treat $t$ as endogenous. This endogeneity is both definitional ($t$ is a function of the dependent variable) and behavioral (factors that are unobserved by the researcher but observed by the assessor may influence $V$ may also influence $T$). *PTHV* develops a model of assessor behavior to identify some instruments and then uses either non-linear two-stage least squares (for equation (8)) or 2SLS (for equation (9)).

Fifth, one must be careful about omitted variable bias, particularly with equation (6). Good data are needed. This is not quite such a big problem with double-sales data, which are used for equations (7) and (9). Even in that case it could arise, however, so data on additions and renovations are desirable.

Note that omitted variable bias is a bigger problem with equation (7) than with (9), because the latter removes all jurisdiction-level variation. In other words, deflating $V$ eliminates the possibility that the estimated impact of a change in $t$ is biased by the omission of other changes at the jurisdiction level, such as changes in public service quality or transportation costs.

Sixth, there is the problem of itemization. If a taxpayer itemizes her deductions, then she gets to deduct property taxes. So a $1 increase in the present value of property taxes does not really cost this taxpayer $1. Perhaps this is reflected in capitalization. In fact, however, mortgage interest payments are also deductible, so an income tax correction applies to both the numerator and denominator of the estimated coefficient, $\beta/r$. If $s$ is the marginal income tax rate, this ratio with full deductibility of interest is $[\beta(1-s)]/[r(1-s)]=\beta/r$. So it is not clear that income taxes need to be considered.

Nevertheless, one might also argue that the denominator is really the opportunity cost of investing in housing, not the mortgage interest rate. This opportunity cost can be thought of as the return on other low-risk, long-term investments, and it is unaffected by deductibility. An alternative view, therefore, is that estimated capitalization should be lower for itemizers than for non-itemizers. For a more detailed treatment of this issue, see deBartolomé and Rosenthal (1999).
Estimates of Property Tax Capitalization

Virtually every study of property tax capitalization finds a statistically significant negative impact of property taxes (or property tax changes) on house values. The vast majority these studies use data from the United States, but a few use data from Canada. Estimates of $\beta$ vary widely, but if $r$ is set at 3 percent, the estimates of $\beta$ for the best studies fall between 15 and 60 percent. See PTHV.

No study yet provides a definitive estimate of $\beta'$, but several recent studies provide some evidence that it is close to 100 percent.

1. Property Taxes and House Values

PTHV is based on double-sales data in Massachusetts for observations before and after revaluation. These revaluations took place in the 1970s. Recall that the $t=m(A/V)$, so when assessment practices are poor, as they were in Massachusetts, effective tax rates vary widely. Revaluation brings effective tax rates into line so there are big changes in effective tax rates, that is, in $\Delta t$ both upward and downward. This study takes advantage of the large variation $\Delta t$ to identify the impact of $t$ on house values.

In Waltham, the community with the best data (i.e. the most complete set of controls for housing characteristics and renovations), the PTHV estimate of $\beta$ equals 21.1% with nonlinear 2SLS (equation (8)) and 22.2% with the 2SLS approximation (equation (9)). These results are based on 353 double-sales observations. Similar results were obtained for 6 other communities.

As noted earlier, one would expect $\beta$ to be 26% if $\beta'$ equals 1.0 and current tax differences are expected to last only 10 years. Because of the legal pressure to revalue in Massachusetts, the 10-year horizon is very reasonable. These estimates correspond to a horizon of 8.5 and 8.0 years, respectively.

As noted earlier, however, these estimates are not quite correct, and need to be divided by the denominator of equation (11). Corrected estimates for Waltham are given in Table 1. Because the correction depends on the effective tax rate, the magnitude of the correction would not be exactly the same in the other communities.
Table 1. Corrected Estimates of Capitalization in Waltham for Property Taxes and House Values

<table>
<thead>
<tr>
<th></th>
<th>Nonlinear Version</th>
<th>Linear Version (equation (9))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate of $b$</td>
<td>7.4233</td>
<td>7.0433</td>
</tr>
<tr>
<td>Value of $I_1^*$</td>
<td>0.0420</td>
<td>0.0420</td>
</tr>
<tr>
<td>Value of $1-bI_1$</td>
<td>0.6882</td>
<td>0.7042</td>
</tr>
<tr>
<td>Original $\beta$</td>
<td>0.2227</td>
<td>0.2113</td>
</tr>
<tr>
<td>Corrected $\beta$</td>
<td><strong>0.3236</strong></td>
<td><strong>0.3000</strong></td>
</tr>
<tr>
<td>Understatement (%)</td>
<td>31.2</td>
<td>29.6</td>
</tr>
<tr>
<td>Implied $N$ (years)</td>
<td>13.2</td>
<td>12.1</td>
</tr>
</tbody>
</table>

Thus, the estimates $\beta$ for Waltham reported in PTHV understate the corrected estimates by about 30 percent. The corrected estimates are consistent with $\beta' = 1$ if people expected pre-revaluation differences in effective tax rates to last for 13.2 and 12.1 years, respectively.

2. Eisenberg

Eisenberg’s 1996 dissertation looks at the capitalization of property taxes in Syracuse before revaluation, when the rates varied considerably. He uses a version of equation (6). He picks up on a point in PTHV (okay, he was my student) that someone borrowing up to her limit to buy a house gets capitalization built into her bid: every increase in property taxes must be offset by a decrease in bid, i.e. in housing price. He also follows up on the point that the capitalization story is cleanest for non-itemizers.

Eisenberg conducted a survey of several hundred residents. He found out if they itemized and if they had borrowed up to their limit. He estimated the degree of capitalization for non-itemizers who were borrowed up to their limit and found that $\beta = 1.0$, exactly 1.0! He also asked households about their expectations concerning revaluation. However, nobody expected Syracuse to revalue (unlike in Massachusetts), so
expectations proved not to matter. deBartolomé and Rosenthal (1999), also account for itemization. They argue that previous studies have not properly specified the role of federal income taxes, but that the resulting errors work in different directions and roughly cancel out on net. Thus, their estimate of capitalization for non-itemizers, about 0.40, is therefore not as high as Eisenberg’s.

Tax Capitalization and Public Policy

Tax capitalization has some bizarre implications for public policy.

If property taxes are fully capitalized, then any tax changes show up in house values immediately and there is no way to escape them. A house with a tax increase must either stay and pay the higher tax or leave and suffer a capital loss. Moreover, the loss is the full present value of the future increases in taxes. People who experience a cut in taxes get a capital gain.

In theory, the gains and losses are effective when a policy change is announced, but in practice there may be some time before information is complete or people really believe that the change will be implemented. The implication for studying capitalization is that the estimated rate of capitalization may be small right after tax changes are announced and then grow to its maximum after implementation.

There is some evidence to support this notion in PTHV. Anticipated property tax changes have a small, but significant impact on house values. However, the degree of capitalization is quite small until tax changes have been fully implemented.

This evidence supports the view that lags in implementation may allow owners to escape some, or even most, of the impact of property tax changes if (a) there is a lag between announcement and implementation and (b) they act right after announcement.

Now consider assessment reform. Assessment reform leads to capital gains for some (those who were overassessed relative to other houses) and to capital losses for others (those who were underassessed). For long-term residents, these changes are fair. A resident who has been underassessed for a long time has been given, in effect, a loan from the city and revaluation just claims back this loan.

But for new residents, these changes are not fair. If someone bought a high-effective-tax house one day and the change is announced the next, this person has a huge capital loss even though she did not benefit
at all from the poor assessment system. In fact, capitalization implies that this person paid a premium for her house.

One way to minimize this fairness problem may be to have a long lag between announcement and implementation. As explained earlier, this lag may allow owners at the time of announcement to escape some of the burden of the tax changes. Moreover, this fairness problem does not arise if houses are revalued upon re-sale, which they were not in Massachusetts or Syracuse.

So a revaluation imposes some unfair gains and losses in order to restore faith in the property tax and local government and to ensure fairness in the future. This trade only makes sense if assessments are updated regularly. Otherwise, small gains and losses are handed out each year as assessment errors mount, which is unfair and undermines the case for reform.

In addition, poor assessments lead to court cases, which the city usually looses. Someone can buy a property cheap because its property taxes are relatively high and then sue the city to get a rebate because of unfairly high taxes. This happened in Boston, to the tune of tens of millions of dollars. The only way to avoid this crazy situation is to keep assessments up to date!

In the case of Proposition 13, assessment growth is fixed at 2%, so the assessment/sales ratio, and hence $t$, declines over time for long-term owners. This cannot be turned into a capital gain because houses are revalued upon sale. But it still represents a gift to owners who stay (and discourages mobility). The U.S. Supreme Court said this was legal, and California voters like the reward to long-term residents. I do not think it is fair.

Another application is to state aid. If a jurisdiction gets aid and cuts its property taxes (we will ignore service impacts for now), this leads to a capital gain to residents. Because aid is sometimes based on property values, there is a strange feedback here: low property values lead to more aid which leads to higher property values which leads to less aid.

As Inman (1978) has pointed out, this leads to problems with a power-equalizing aid program. This type of program gives a higher matching rate to lower-wealth jurisdictions. This higher matching rate raises property values in lower-wealth jurisdictions and therefore leads to lower aid for them as time goes on. In other words, the property-value impacts undermine the objectives of the aid program. As I show in a paper I am
working on, however, this feedback effect does not arise with another important type of aid called foundation aid. Foundation aid provides higher block grants (not higher matching rates) to lower-wealth jurisdictions, but with a fixed budget for the program, the amount of aid they receive is not altered by property value effects, should they arise.
References


