

THE SPATIAL DISTRIBUTION OF HOUSING-RELATED TAX BENEFITS
IN THE UNITED STATES

Joseph Gyourko
Todd Sinai

April 3, 2001

This research was supported by the Brookings Institution's Center on Urban and Metropolitan Policy and the Research Sponsor Program of the Zell/Lurie Real Estate Center at The Wharton School, University of Pennsylvania. Seminar participants at the Brookings Institution and the Wharton School provided helpful suggestions. In addition, Bill Dickens, Tony Downs, Bill Gale, and James Poterba provided thoughtful comments. Finally, we thank Kiwan Lee for providing excellent research assistance and Paul Amos for the GIS plots.

© 2001 by Todd Sinai and Joseph Gyourko. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

ABSTRACT

Using 1990 Census tract-level data, we estimate how tax subsidies to owner-occupied housing are distributed spatially across the United States, calculating their value as the difference in taxes currently paid by home owners and the taxes owners would pay if there were no preference for investing in one's home relative to other assets. The \$164 billion national tax subsidy is highly skewed spatially with a few areas receiving large subsidies and most areas receiving small ones. If the program were self-financed on a lump sum basis, less than 20 percent of states and 10 percent of metropolitan areas would have net positive subsidies. These few metropolitan areas are situated almost exclusively along the California coast and in the Northeast from Washington, DC to Boston. At the state level, California stands out because it receives 25 percent of the national aggregate subsidy flow while being home to only 10 percent of the country's owners. At the metropolitan area level, owners in just three large CMSAs receive over 75 percent of all positive net benefits.

Todd Sinai
Wharton School
University of Pennsylvania
308 Lauder-Fischer Hall
256 South 37th Street
Philadelphia, PA 19104-6330
(215) 898-5390
and NBER
sinai@wharton.upenn.edu

Joseph Gyourko
Wharton School
University of Pennsylvania
313 Lauder-Fischer Hall
256 South 37th Street
Philadelphia, PA 19104-6330
(215) 898-3003
gyourko@wharton.upenn.edu

Introduction

With 65 percent of U.S. households owning their own homes at a given point in time and a higher percentage owning a house at some point during their lifetimes, the tax treatment of owner-occupied housing is one aspect of the tax code that affects many people's daily lives. It is not surprising that it is well studied in many aspects. The mortgage interest and property tax deductions, in conjunction with the non-taxation of imputed rent, reduce the cost of owner-occupied housing relative to other investments [Hendershott and Slemrod (1983), Poterba (1984)], encourage home ownership and higher housing consumption [Rosen (1979), King (1980), Henderson and Ioannides (1989)], and perhaps even lead to overinvestment in the asset class [Mills (1987), Feldstein (1987)]. The subsidy may raise house prices [Capozza, Green, and Hendershott (1996), Bruce and Holtz-Eakin (1998), Sinai (1998)] and encourage suburbanization [Gyourko and Voith (2001)]. The tax treatment of owner-occupied housing also favors high-income people or those who own expensive houses [Poterba (1992)], and its potentially substantial contribution to families' net worth has proven to be a political sticking point to eliminating it in any flat tax proposal.

Despite the considerable attention paid to the tax treatment of owner-occupied housing, little is known about the geographical distribution of this subsidy. However, since housing markets are inextricably tied to a physical location, knowing how the tax subsidy varies spatially is useful for helping determine the effects of a change in the tax code. At the least, it is important to consider the extent to which some areas of the country receive a greater share of the annual flow of tax code-related benefits to owner-occupied housing, both to see if resources are flowing from some locations to others and to determine if some areas might be more sensitive to a policy

change. In addition, the spatial distribution may provide some insight into the political economy of changing that tax provision.

In this paper, we take a first step in answering the larger question of what the economic impact of changing the tax treatment of owner-occupied housing would be by documenting where the tax benefits flow spatially, both within and across states and within and across metropolitan areas. We calculate the value of the tax subsidy to owner-occupied housing as the difference in taxes currently paid by home owners and the taxes they would pay if the tax code treated them like landlords. Unlike the current code, such tax treatment would not provide a preference for investing in one's home relative to other assets.

While our focus is on how much the tax subsidy to owner-occupied housing would decline in some areas relative to others if neutral taxation of owner-occupied housing were introduced, we also discuss the two extremes in terms of the possible ranges and magnitudes of the effect on house prices or user costs of owning. However, we do not measure the redistributive effects of other tax benefits, tax expenditures, or the progressive income tax. While these other aspects of the tax code are certain to redistribute income spatially in ways that may offset or augment the effects of the tax subsidy to home ownership measured here, we believe it is most appropriate to view the effects of each feature of the tax code in isolation. We also do not consider general equilibrium effects associated with the possibility that reducing the subsidy to owner-occupied housing may lead to a decline in house values. This outcome would reduce the tax subsidy to owner-occupied housing further since the opportunity cost of equity would be lower.

Using 1990 Census tract-level data, we estimate the tax subsidy to owner-occupied

housing for the nation as a whole to be quite large, almost \$164 billion in 1989, corresponding to \$2,802 per home owner and \$1,815 per household.

The subsidy is highly skewed spatially. In only 12 cases does a state's share of the aggregate subsidy exceed its share of the nation's owner-occupied housing units. California alone reaps \$41.5 billion, or 25 percent, of the gross benefits under the program while being home to only 10 percent of the owner-occupied units in the country. The program also effectively transfers just over \$18 billion from census tracts in cities to those outside cities.

Across metropolitan areas, only ten percent of them receive more than the national average subsidy per owner-occupier. Most of these areas lie along the California coast and Amtrak's Northeast Corridor running from Washington, DC, to Boston, MA. Even among this small group of areas, benefit flows are highly concentrated. When program costs are accounted for by assuming the program is self-financed via lump sum payments made by each household, we find that owners in just three large Consolidated Metropolitan Statistical Areas (CMSAs) – Los Angeles-Riverside-Orange County, New York-Northern New Jersey, and San Francisco-Oakland-San Jose – receive just over 75 percent of all positive net benefit flows measured at the metropolitan area level.

The spatial distribution of benefits within metropolitan areas can be skewed as well, although this varies considerably across areas. In many metropolitan areas, especially the smaller ones away from the two coasts, benefits tend to be relatively evenly distributed across owners. In contrast, in a number of the larger metropolitan areas, the top quarter of owners receives 70 percent or more of the total subsidy flowing to the area.

The economic implications of these results for any change in the current tax code depend

upon whether or not the subsidy is capitalized into land prices. If it is fully capitalized, eliminating the subsidy will not affect the user cost of owning, but there would be significant wealth effects for many owners. Back-of-the-envelope calculations suggest that the perpetuity value of the subsidy amounts to about a fifth of property value on average in the United States. While we presume the savings associated with eliminating the subsidy would be redistributed in lump-sum fashion, the net wealth effect still is likely to be significant in many areas. For example, there are twenty metropolitan areas, including many densely populated ones centered around Boston, New York City, Washington, D.C., Los Angeles, and San Francisco, for which the present value of the subsidy flow is greater than 25 percent of house values.

If the elimination of the tax subsidy is not capitalized into land prices, the user cost of ownership must rise. Our calculations suggest that the increase would amount to between 3 and 5 percent of annual household income in about three-quarters of the nation's metropolitan areas. However, in about 10 percent of metropolitan areas, including many in California, Hawaii, and Massachusetts, the rise in user costs would amount to 10 percent of income or more. On the other hand, our analysis indicates that returning the savings from eliminating the subsidy as a lump-sum refundable credit would more than offset the user cost increase in most areas, even assuming households do not reduce their housing consumption in response.

Finally, it is noteworthy that our findings are not simply a reflection of the progressivity of the tax code. One would expect benefits to be skewed towards high income households under a progressive rate structure, and if these households live close to each other that could account for the spatial patterns we identify. However, high income, high tax bracket owners tend to reside in disproportionately valuable homes so that the tax code is interacting with housing

consumption to provide an extremely skewed subsidy distribution. Stated differently, while high income owners certainly do pay a large share of taxes, they receive an even larger share of this program's benefit flows.

The remainder of the paper proceeds as follows. In the first section, we describe the tax subsidy to owner-occupied housing and how we measure it. Section two reports our results, beginning with an analysis of how this tax program redistributes income across states and moving through progressively smaller geographies before concluding with a description of the spatial distribution of the tax benefits within some example metropolitan areas. Section three then considers the economic implications of our findings in terms of potential capitalization and/or user cost effects. Section four provides an analysis of the factors generating the spatial effects that we identify. Finally, there is a brief conclusion and summary.

I. Measuring Housing-Related Tax Benefits

Determining the Subsidy Under the Current System

In a perfect housing asset market, the marginal home owner will invest in owner-occupied housing until the point where the annual cost she incurs exactly equals the rent she would have to pay as a tenant in the same property [Hendershott and Slemrod (1983), Poterba (1984)]. That user cost is described in equation (1) and takes into account the fact that implicit rental income is untaxed while mortgage interest and property taxes are deductible for itemizers:

$$(1) \quad R_H = (1-\tau_{\text{ded}})\alpha i + (1-\tau_{\text{ded}})\tau_p + (1-\tau_{\text{int}})(1-\alpha)r + M + \delta - \Pi^H.$$

The left-hand side variable, R_H , is the implicit rent per dollar of housing value the owner pays herself. In equilibrium, it must equal the cost of owner occupancy per dollar of housing value.

These costs include: (a) the after-tax cost of mortgage interest, $(1-\tau_{\text{ded}})\alpha i$, where α is the loan-to-value ratio on the house, i is the mortgage interest rate, and τ_{ded} is the owner-occupier's marginal tax rate which equals her marginal rate (denoted τ_{int}) if she itemizes and equals zero otherwise; (b) the after-tax cost of property tax payments, $(1-\tau_{\text{ded}})\tau_p$, with τ_p the effective property tax rate; (c) the after-tax opportunity cost of investing equity in the house rather than in some other investment at rate of return, r ; this is given by $(1-\tau_{\text{int}})(1-\alpha)r$ and is a cost to all owners, whether they itemize or not;¹ (d) annual maintenance costs per unit of housing which are given by M ; (e) the cost of true economic depreciation per unit of house which is assumed to occur at rate δ ; and (f) any annual appreciation in the house value, Π^H , which reduces the carry cost.²

The components of the subsidy to owner-occupied housing can be highlighted by comparing the current tax treatment to that under a different tax code which does not favor owner-occupied housing. Treating the home owner like a landlord, for example, taxes the house like any other asset. Since the home owner essentially rents her house to herself for R_H , neutral tax treatment would require taxing the implicit rental income on one's home. If treated like landlords, owner occupiers also would be able to deduct maintenance expenses and depreciation in addition to the mortgage interest and local property taxes presently allowed. Making these adjustments (and assuming accrual taxation on capital gains for clarity of exposition) would yield

¹This notation assumes that the opportunity cost of tying up equity in a house is foregoing taxable returns. If the home owner were to invest in a tax-exempt asset instead, we assume the return would be $(1-\vartheta)r$ rather than r , yielding the same after-tax return. To the extent that the home owner has a lower tax rate than the marginal investor in municipal bonds, her opportunity cost of equity would be less than $(1-\vartheta)r$.

²This specification treats capital gains on housing as untaxed and realized every year. Given that there now is a \$250,000 capital gains exclusion (\$500,000 for married couples filing jointly) that can be applied every other year, this is not unrealistic. Even in 1989, the assumption of no capital gains taxation on housing was valid for the vast majority of households.

a different perfectly competitive equilibrium rent given in equation (2):

$$(2) \quad R_H' = (1-\tau)\alpha i + (1-\tau)\tau_p + (1-\tau)(1-\alpha)r + \tau R_H' + (1-\tau)M + (1-\tau)\delta - (1-\tau)\Pi^H.$$

where $\tau R_H'$ is the tax due on imputed rent. Grouping the R_H' terms and dividing both sides by $(1-\tau)$, we obtain:

$$(3) \quad R_H' = \alpha i + \tau_p + (1-\alpha)r + M + \delta - \Pi^H.$$

Comparing equations (1) and (2) demonstrates that three factors determine the differences in the user costs under the two tax systems. User costs of owning would be higher by the value of the untaxed imputed rent if tax neutrality were imposed, but the difference would be reduced by the values of the maintenance and depreciation deductions. More formally, the difference in user costs under the two tax regimes, $R_H' - R_H$, represents the subsidy to owner occupancy under the current code and is captured in equation (4):

$$(4) \quad R_H' - R_H = \tau_{ded}\alpha i + \tau_{ded}(\tau_p) + \tau_{int}(1-\alpha)r.$$

This equation measures the change in user costs, or implicit rent for owner-occupiers, that would result if the current tax system were modified so that one dollar of investment in owner-occupied housing was not favored over other assets.³ This is the tax subsidy to owner-occupied housing and it can be decomposed into three factors: (a) the tax value of home mortgage interest deductions ($\tau_{ded}*\alpha*i$); (b) the tax value of local property tax deductions ($\tau_{ded}*\tau_p$); and (c) the tax that would have been paid on the equity invested in the home had it been invested elsewhere

³We have conveniently abstracted from *how many* housing dollars on which a home owning family receives a subsidy. A change in the tax treatment of owner-occupied housing might affect house values. Since we measure the subsidy on a per dollar basis, we exclude the possibility that there is a second order effect through changes in house prices. This is done for two reasons. First, as noted in the Introduction, determining exactly how a change in the subsidy would be capitalized into house values is beyond the scope of this paper. Second, any change in house price would only increase the magnitudes of our estimates. For example, if the benefit to owner-occupied housing were reduced, house prices might also fall, further decreasing the subsidy.

$(\tau_{\text{int}}*(1-\alpha)*r)$. By estimating these components, we can determine the total subsidy to owner-occupied housing under the current tax code.⁴

Data and Estimation Strategy

Census tract level information in the STF3 files of the 1990 decennial census for all fifty states plus the District of Columbia are our primary data. We use census data rather than tax return data because the census reports geocodes down to the tract level while we can only obtain state identifiers with the tax data. Tract codes enable us not only to identify center city versus suburban areas, but to identify specific communities in the suburbs. The added location detail is critical when examining finer spatial distributions within an individual metropolitan area. And, the census data report most of the crucial elements – including the distributions of house value and family income, along with certain demographics – needed to compute reasonably accurate estimates of the tax benefit.⁵

We start by computing the distribution of household income among home owners at the tract level.⁶ For each tract, we divide the household income distribution into deciles and assign

⁴We do not intend to imply that the mortgage interest or local property tax deductions themselves create subsidies to owner occupiers. The subsidy arises from the non-taxation of imputed rent in conjunction with those deductions. However, mathematically the subsidy can be represented by the three terms in equation (4).

⁵Another possible data source is the *American Housing Survey* (AHS). Unlike the Census tract-level data, the AHS reports income and house value for individual houses. Unfortunately, the national files of the AHS do not contain state-level identifiers and the metropolitan area files only label central city status, making the *Survey* a poor choice for our application.

⁶All tax benefit figures reported in this paper are based on tract-level data that aggregates household income across its various sources. Tract-level data also are available on seven components of income (wages and salary, interest and dividends, social security, public assistance, farm and non-farm self employment income, retirement income, and “other”). Experimentation showed that our estimates of the value of tax benefits are not sensitive to whether aggregate or disaggregate income data are employed. In both series, capital gains information is missing. Hence, the data almost certainly under report income at the upper end of the distribution. We do not believe this is an especially serious problem for 1989 because the top tax bracket is reached before one gets too far up the income distribution. Thus, we probably do not underestimate the tax bracket of most of the households for which income is most likely under-reported.

the median income for each decile to all the households in that category. Thus, the lowest-income one-tenth of the households are assumed to have an income equal to that of the fifth percentile for the tract, the next lowest-income tenth of the households are assigned an income equal to that of the 15th percentile for the tract, and so forth.

We then map tract-level information on the distribution of house values to incomes by assigning to households in each decile of the income distribution the value corresponding to the same decile of the house value distribution. For example, we assume that the household in the 5th percentile of the income distribution for the tract also owns the home in the 5th percentile of the housing price distribution for the same tract.⁷

We begin estimating the components of equation (4) with each tract-decile's weighted average value of the mortgage interest deduction, computing the tax value as the difference in tax bills with and without it. The mortgage interest deduction itself is defined as $P_H * \alpha * i$. Leverage ratios, α , were allowed to vary by age based on data computed from the 1989 *Survey of Consumer Finances (SCF)*. A weighted average leverage for each tract was computed based on the tract's age distribution.⁸ The mortgage interest rate, i , was calculated by taking an average

⁷This matching process presumes that owners and renters in a tract have identical income distributions. Fortunately, our results are robust to alternative assumptions. For example, if we assume an extreme case in which all the owners in a tract have a higher income than any of the renters, and houses are matched to owners so that the highest income owner owns the highest value house, the next highest income owner occupies the next highest valued house, and so forth, none of our spatial results at the national, state, or metropolitan levels change in a material way. The spatial distributions are only slightly more skewed than those reported below in the text, though in this case, the estimated aggregate subsidy is about 25 percent higher. This latter result is to be expected given that in this matching scheme owners always have the highest tax rates in each tract since they have the highest incomes.

⁸Loan-to-value ratios by age are as follows: 20-24 year olds – 53.6 percent; 25-29 year olds – 70.2 percent; 30-34 year olds – 66.1 percent; 35-39 year olds – 54.7 percent; 40-44 year olds – 48.9 percent; 45-49 year olds – 41.5 percent; 50-54 year olds – 31.0 percent; 55-59 year olds – 24.0 percent; 60-64 year olds – 20.4 percent; 65-69 year olds – 10.3 percent; 70-74 year olds – 10.3 percent; and 75+ year olds – 2.4 percent. In addition, our findings are not sensitive to assuming 35 percent leverage, the national average in 1989 according to *American Housing Survey* data, for all tracts.

from the 1989 *SCF*, and that rate was equal to 9.84 percent.

The tax value of the mortgage interest deduction can differ from mortgage interest paid times the marginal tax rate for three reasons. First, only families that itemize on their tax returns receive any benefit on the margin from the deductibility of mortgage interest. Also, only the excess of the mortgage interest deduction plus other itemized deductions over the standard deduction has value for a taxpayer. Therefore, we would only multiply the portion of mortgage interest in excess of the standard deduction (after itemizing all other non-housing related deductions first) by the tax rate. Additionally, since the tax schedule is nonlinear, taking the mortgage interest deduction may lower the taxpayer's marginal and average tax rates.

The actual value of the tax benefits also depends on certain demographic data that are likely to affect the number of exemptions and the overall amount of deductions. Tract level data that are available include the distribution of whether households are single, married, or single with children; the percentage of households with children; and the percentage of households over 65 years of age.

Unfortunately, the census data lack information on most non-housing categories of potential tax deductions. We compute mortgage interest, state, and property tax deductions, but we do not observe medical expenses, charitable giving, deductible interest (other than for a home mortgage), and several other miscellaneous categories. Charitable giving alone accounted for \$55 billion of the \$432 billion of total itemized deductions reported on tax returns in 1989, placing it behind only mortgage interest (\$169 billion) and state and local taxes (\$81 billion). Thus, these omissions potentially are severe.

Two counterbalancing problems arise from underestimating possible deductions. First,

we would be more likely to incorrectly assume the family does not itemize. This error would cause us to underestimate the tax value of the mortgage interest and property tax deductions since less would be deducted at the margin. On the other hand, undercounting deductions for itemizers could increase the tax value we do measure since the remaining deductions are applied against higher marginal tax rates.

Consequently, we impute missing tax deductions to our census data based on data from the Department of the Treasury's *Statistics of Income (SOI)* public use tax micro sample. A modified Heckman-style sample selection model is employed to correct for the selective observing of deductions only by itemizers, with the details reported in the appendix.⁹ Following this imputation, federal and state tax rates and implied tax benefit amounts are computed using the National Bureau of Economic Research's (NBER) TaxSim program.

The second component of equation (4) involves the value of the deduction of local property taxes. Property tax payments themselves are defined as $P_H * \tau_p$, where τ_p is the average effective property tax rate. This is allowed to vary by state using data from the Advisory Commission on Intergovernmental Relations (ACIR (1987)).¹⁰ The tax value of the deduction associated with these tax payments then is computed the same way as for the mortgage interest deduction.

⁹The imputation results indicate that underestimation of deductions and therefore underestimation of itemizers was the biggest problem associated with sample selection. This turns out to be important because the underestimation of itemizers was not random across space. In high house value and high income tax states such as California, not observing non-housing deductions only infrequently caused us to miscategorize an owner family as a non-itemizer. Home mortgage interest, local property taxes, and state income taxes generally were sufficient to make California residents itemizers. This was not the case in many states with lower house values and lower state taxes. Hence, the imputation has an important effect on the measured spatial distribution of program benefits.

¹⁰ The ACIR did not report state-by-state breakdowns for 1989, so we use the 1987 data. We have also experimented with assuming a 1 percent and 1.5 percent national average effective rate. Our findings are not sensitive to these changes.

The last component in equation (4) arises from the fact that the government does not tax as income the implicit return on equity an owner invests in the home. Precisely what home owners would have done with their money had they not purchased a home cannot be known with certainty, of course. We assume that for most owners the alternative is a relatively safe investment with a duration similar to mortgages. Consequently, we use the 8.71 percent yield on seven year Treasuries in 1989 to represent the foregone interest that could have been earned on home equity.¹¹

The value of the non-taxation of the return on equity invested in housing is computed in two steps. First, we calculate the opportunity cost of the equity in one's home, or $P_H \cdot (1 - \alpha) \cdot r$, where r is the 8.71 percent Treasury yield. We then calculate the difference in tax liabilities between the cases in which the family invested the home equity in taxable form and in which they held untaxed housing. This approach accounts for the possibility that a family might move into a higher marginal tax bracket if the return on its housing equity was taxed.

The procedure for estimating equation (4) is represented graphically in Figure 1 which shows a tax schedule with three marginal tax brackets. A home-owning family with no housing-related deductions would have a taxable income (TI) of Y_1 . However, if they were not owners, they may have invested their housing equity in a vehicle that yielded a taxable return which would raise their TI to Y_2 . Thus, Y_2 is the counterfactual TI for a home-owning family if it were to stop being an owner. Starting with that TI, we can compute the tax value of each of the three aforementioned deductions. With an taxable income of Y_2 , this hypothetical family would have

¹¹One extreme would be to assume that the equity portion of the house value would have been put under the home owner's mattress, earning a zero nominal return. In that case, user costs are lower under the current tax code only because of the value associated with mortgage interest and local property tax deductions.

a tax liability of T_1 . Assume that claiming the home mortgage interest deduction (HMI) would lower TI to $Y_2 - \text{HMI}$ (presuming for simplicity that all of HMI was above the standard deduction) and the tax liability to T_2 . Therefore, the tax savings for this family from the mortgage interest deduction is $T_1 - T_2$.

Although in this example the mortgage interest deduction does not move the family into a lower tax bracket, the property tax deduction does. Beginning with TI equal to $Y_2 - \text{HMI}$, we can compute the tax savings from the property tax deduction as the tax bill with only the mortgage interest deduction, T_2 , minus the tax bill with both the mortgage interest and property tax deductions, T_3 . In this case, T_2 and T_3 span a kink in the tax schedule, but still account for the fact that the average tax rate is less than the marginal tax rate at $Y_2 - \text{HMI}$.

Finally, we compute the value of the non-taxation of the return on housing equity. Because the return on housing equity is not included in TI, taxable income is measured at Y_1 instead of the greater amount Y_2 . The tax value of not including that income is measured as the change in tax between T_3 (the tax bill corresponding to an TI of $Y_2 - \text{HMI} - T_p$) and T_4 (the tax bill corresponding to an TI of $Y_1 - \text{HMI} - T_p$). If there is no opportunity cost for the housing equity, $Y_1 = Y_2$ and the tax savings is zero.

It is apparent from Figure 1 that the order in which the deductions are taken matters when the tax schedule is not linear. For example, $T_1 - T_2 > T_3 - T_4$, even though $\text{HMI} < Y_1 - Y_2$. After adding back the implicit return on housing equity, we compute the deductions in the following order: (a) tax savings from the mortgage interest deduction; (b) the tax savings associated with the property tax deduction; and (c) the savings from the return on housing equity being untaxed. We have repeated the estimation using all six possible sequences in which the deductions can be

taken. While the relative magnitudes of the categories do change, the differences are minor.

II. Results

Summary Statistics

There were 90.2 million households in the nation in 1989, with over 61 million residing outside of central cities.¹² The propensity to own is high, with 58.4 million or 64 percent of all households being owner-occupiers – 71 percent of those living outside of central cities and 50 percent of households living inside of central cities.

The value of tax benefits to owner occupiers is great. Table 1 reports aggregate, per owner-occupied housing unit, and per household values of the aggregate tax benefit for the U.S. The gross value to owners of all housing-related tax benefits for the country in 1989 was nearly \$164 billion (top panel). Of this total, 62 percent derives from the untaxed return on home equity. Nearly \$43 billion, or 26 percent of total tax benefits, is due to mortgage interest.¹³ This fraction is lower than the approximately 35 percent of total house value that is debt financed because not all available mortgage interest deductions are in excess of the standard deduction. The remaining \$20 billion in housing-related tax benefits is generated from the deduction of local property taxes.

Gross program benefits per owner-occupied housing unit are \$2,802 (middle panel). This

¹²This calculation is made based on central city designations of the Office of Management and Budget.

¹³Our \$43 billion estimate of the value of mortgage interest deductions overstates the value based on computations using the SOI by about \$6 billion. One reason is that the SOI calculation will underestimate the true subsidy value since it does not add back foregone equity as income. In addition, we suspect that some of the discrepancy is due to the fact that our deduction imputation procedure does not take account of the possibility that taxpayers who itemize tend to have deductions in multiple categories. Hence, we probably underestimate the total amount of deductions and therefore apply a higher tax rate to housing deductions than we do with the SOI, where we observe each taxpayer's actual deductions.

figure results from dividing the \$163.8 billion in aggregate benefits by the 58.4 million owned units nationally. The \$1,815 in tax benefits per household represents the cost per household needed to fund the program if it were self-financed. Under that assumption, net benefits to owners on average are only \$987 (\$2,802-\$1,815). Renters suffer a net outflow in the amount of the \$1,815 mean program cost since their tax treatment is the same under the current code and one that taxes owner-occupiers like landlords.¹⁴

Incorporating program costs into the analysis further permits us to compute at the national level the extent to which the program allocates resources to tracts in central cities versus those in outlying areas. For ease of exposition, we refer to any census tract not in a central city as being in the suburbs.¹⁵ Differences in subsidy flows between city and suburban tracts are measured using the variable we term the Suburban-City Benefit Gap (SCGAP). This is computed as follows for each census tract:

$$\text{SCGAP} = (\text{Aggregate Suburban Tax Benefit Value} - \text{Aggregate City Tax Benefit Value}) - (\$1,815 * [\text{Suburban Households} - \text{City Households}])$$

This is the difference in the value of aggregate tax benefits realized in each area, adjusted for the average program cost that each household is presumed to pay on a lump sum basis. Overall, the

¹⁴One might argue that the current code subsidizes renters because landlords are able to deduct various expenses and competition may force them to pass along some of the tax code-related benefit to their tenants. However, comparing the current tax system to our neutral one nets out any subsidy to renters. Since the taxation of landlords is unchanged across the two tax systems, eliminating the subsidy for owner-occupiers will not affect renters, other than by saving them the \$1,815 mean program cost. The current code would subsidize renters if landlords are allowed to depreciate income properties for tax purposes faster than true economic depreciation – and if that were passed along to renters in the form of lower rent. However, taxing owner-occupiers like landlords would not change the depreciation schedule and, therefore, would not affect the subsidy to renters. We would overestimate the loss to *owner-occupiers* of “eliminating” their subsidy since it would be replaced by another one – accelerated depreciation. However, while one could argue that the statutory depreciable life in 1981 (of 15 years) was shorter than true economic depreciation, the situation post-1986 is more akin to one in which economic depreciation is not very different from statutory depreciation.

¹⁵We have performed the analysis restricting the data to tracts in metropolitan areas so as to cut down on the number of truly rural observations. None of our findings are materially affected by this change in sample.

tax preference for owner-occupied housing transfers resources from central city areas to outlying areas. Nationally, the SCGAP calculation shows a net transfer of \$18.2 billion from central cities to suburbs, amounting to 11 percent of the gross aggregate subsidy.

State-Level Results

Table 2 presents data on gross and net program benefits at the state level. The first column reports the value of total tax code-related benefits per owner-occupied housing unit for the fifty states and the District of Columbia. There is wide variation around the \$2,802 average value for the nation reported in Table 1. The state means range from a low of \$775 in South Dakota to \$9,181 in Hawaii.

One way of characterizing the state-to-state differences in benefits is by comparing the state's share of the gross aggregate tax benefits to its share of the country's owners.¹⁶ Column 2 of Table 2 reports this ratio. Only 11 states plus the District of Columbia have subsidy ratios greater than one. California is a prime example, having 10 percent of the country's owned units and receiving more than 25 percent of the country's aggregate tax code-related benefits to owners, for a subsidy ratio of 2.6. Hawaii has the highest subsidy ratio, with a tax benefit share of 1.1 percent but only 0.3 percent of owned units. South Dakota has the lowest subsidy ratio, as its subsidy share is only one quarter that of its share of the nation's home owners.

These tax benefits have to be paid for in some way, and if we assume the program is self funding, the lump-sum cost per household must equal the \$1,815 national average benefit per household. Column 3 Table 2 reports net benefit figures per owned unit, with this figure equaling the total tax benefit per owned unit from the first column less \$1,815. The negative

¹⁶ The aggregate value of gross tax benefits simply is the product of gross benefits per owner in each state times the total number of owners in the state.

values reported for 26 states indicate that the per household lump-sum program cost exceeded the average tax benefit in these areas. However, there is not much variation in the net negative benefits per owned unit among those states, and the biggest negative transfer states do not have large populations.

Net program benefits in the state, which are the sum of net benefits to owners and renters, are reported in Column 4. As discussed above, each renter household in each state is presumed to pay \$1,815 to support the program. A positive number indicates the state receives a net transfer from other states under the program, assuming it is lump-sum financed. There are only twelve states (including the District of Columbia) who are net recipients under the program once renter costs are taken into account. Figure 2 plots this state-level net transfer series and highlights how skewed the benefits are even among these dozen areas. California is the biggest recipient in aggregate, receiving over \$22 billion from the rest of the nation -- more than all the other net positive beneficiaries combined. Even given California's large population, this amounts to \$2,211 per household and \$3,953 per owner-occupied unit. Owners in Hawaii receive much bigger transfers of \$5,994 on average, but the smaller number of owners puts the state's net benefit at only \$1.1 billion. To put these numbers in perspective, the mean annual benefit paid to poor families nationwide on AFDC in 1990 was \$4,468 according to the 1998 *Green Book Overview of Entitlement Programs*.

Texas clearly is the biggest loser on a statewide basis, suffering a negative net transfer of \$6 billion. This amounts to \$997 for each of the approximately six million households in the state. Other large aggregate losers include the high population states of Florida, Ohio, Pennsylvania, and Michigan.

Between-Metropolitan Area Results

Examining the distribution of subsidy across metropolitan areas further highlights how the spatial skewness of the program benefits increases as we move to more disaggregated geographies. Table 3 reports data analogous to that in Table 2, this time for metropolitan areas rather than states. All 262 metropolitan areas that were considered Metropolitan Statistical Areas (MSAs) or Consolidated Metropolitan Statistical Areas (CMSAs) in 1990 are included. In addition, for each state, we aggregated the census tracts that were not in government-defined MSAs or CMSAs and defined them as their own ‘area’ so that there are 312 total areas listed in the table. For example, in Alaska these tracts are termed the ‘Non-MSA Alaska’ area. Including the tracts in these areas allows us to see the distribution of subsidy inside and outside of metropolitan areas.

The information in column three on the Subsidy Ratio provides the first indication that the subsidy distribution is even more skewed when measured at the metropolitan area level. There are only 30 metropolitan areas, or approximately 10 percent, with Subsidy Ratios in excess of one. This compares to the 12 (or 20 percent of) states, including the District of Columbia, with ratios above one. Of the 30 areas, eight are in California, eight are in Massachusetts, Rhode Island, Connecticut, or New Hampshire and centered around the Boston area, two are in Hawaii, with the others scattered across the country in large population centers such as New York City, Washington, Chicago, Seattle, Philadelphia, and Atlanta. These 30 metropolitan areas are relatively populous, containing almost 30 percent of the nation’s owner-occupied units.

Even when one scales the data to look at benefits per owned unit, there is a consistent pattern of highly spatially skewed subsidies at the metropolitan area level. A relatively small, but

not minuscule, fraction of owners in a few areas are doing very well under the current tax code, with a host of owners in the vast majority of metropolitan areas having benefit flow levels fairly close to the mean program cost. Column 3 reports the value of net tax benefits per owned unit, which net out the mean program cost of \$1,815 from gross program benefits per owner. Ninety of the 312 areas have positive net benefit values for owners and contain roughly half of all owners nationwide. Focusing on those with gross benefits of at least \$3,600, double the mean program cost, finds 19 areas that meet the criterion. There are over 12.3 million owners in these areas, which amounts to 21 percent of all owners throughout the nation.

Figure 3's plot of the aggregate net transfer data from column 4 of Table 3 highlights just how spatially concentrated are overall program benefits. After netting out all program costs, including those paid by renters, only 28 of the 312 metropolitan areas have positive aggregate net benefits. Only five of these areas receive strictly more than \$1 billion per year in net benefit (although Honolulu is very close). Three CMSAs alone – San Francisco-Oakland-San Jose, Los Angeles-Riverside-Orange County, and New York City-New Jersey – receive \$36.5 billion of the total \$47.7 billion in positive net transfers nationwide. These three CMSAs are densely populated, containing 14 percent of the nation's owners and 16 percent of the nation's households. However, the figure makes clear just how tightly spatially targeted are program benefits across metropolitan areas.¹⁷

Within-Metropolitan Area Results

We now ask whether one sees the same kind of spatial skewness in the value of the housing subsidy when we focus within metropolitan areas. It turns out that in some metropolitan

¹⁷ Scaling net transfers by the number of households does reduce the skewness somewhat. However, the bulk of areas still experience small to modest net negative benefit flows and very few areas receive large positive

areas the distribution of the housing subsidy is fairly equitable while in others, including many of the largest cities, a small portion of the population captures the bulk of the tax benefits. Figure 4 reports plots of the cumulative aggregate subsidy against the cumulative percentage of owners for a small sample of metropolitan areas. This chart allows one to determine visually what share of the subsidy flow to an area is captured by any given fraction of owners in the region. If the plot for a metropolitan area were the a straight line through the origin, then 25 percent of the owners in the area would receive 25 percent of the subsidy flowing to the area, 50 percent of the owners would receive half the subsidy, and so on. The more outwardly bowed the plot, the more concentrated the region's subsidy flow is among a relatively few owners. The 45-degree line is plotted as a reference and the thick curved line is the actual cumulative distribution of subsidy in the MSA.

Figure 4 shows stark differences across metropolitan areas. While no region's plot is a straight line, those for Appleton, WI, and Madison, WI, are closest. In communities like these, the subsidy is not highly concentrated among a relatively few owners. The picture is quite different in other metropolitan areas such as Atlanta, Chicago, Dallas, Houston, and Philadelphia. In these areas, typically 40 percent or more of all subsidies flowing to the area accrue to no more than 25 percent of the owners in the metropolitan area. Generally, the less populous metropolitan areas off the two coasts have the least skewed subsidy distributions. Coastal areas and the more populous metropolitan areas tend to have more skewed distributions in the sense that a relatively few owners capture most of the benefits in the region.¹⁸

transfers.

¹⁸ We also looked for spatial skewness within metropolitan areas. For a small sample of areas, we mapped tax benefits per household and examined whether the high benefit households tended to live close to one another. Overall, the high benefit tracts tended to clump together, but there are relatively smooth gradients from low benefit

III. Implications: Capitalization and User Costs

The impact of any change in the tax treatment of owner-occupied housing on households depends upon the extent to which the subsidy is capitalized into prices. In this section, we discuss two extreme cases -- the subsidy being fully capitalized into house prices and no capitalization whatsoever -- to illustrate the range of potential outcomes.

To help gauge how large these impacts might be, we created two variables for each tract in each metropolitan area. CAPSUB provides an upper bound on the possible size of capitalization relative to property value and is defined as the ratio of the perpetuity value of the subsidy (PVSUB) to tract average house value (AVGHV). INCSUB represents how much user costs could change relative to income if there were no capitalization whatsoever and is computed as the ratio of the annual subsidy for 1989 (SUBSIDY) to tract average household income (AVGINC).

The numerator of CAPSUB represents the present value of the annual tax subsidies assuming they are unchanged in perpetuity. The perpetuity value for any tract i in metropolitan area j is defined as the 1989 subsidy value for the tract divided by the capitalization rate in the metropolitan area ($CAPRATE_j$), which is defined as the rent to price ratio.¹⁹ Since rents equal user costs in equilibrium, we employ the user cost divided by house value for each tract as a proxy for the cap rate and take the owner-weighted mean across all tracts in the metropolitan area.²⁰ The denominator of CAPSUB is the tract average house value. Thus,

to high benefit areas.

¹⁹The capitalization rate used here is based on a real estate concept meant to reflect the ratio of the net operating income on a property to its asset value. For those more familiar with financial multiples, it is akin to the inverse of the price-earnings ratio on a stock.

²⁰To better understand this, it is useful to rewrite Equation (1) more simply as USER

$$\text{CAPSUB}_{ij} = \text{PVSUB}_{ij} / \text{AVGHV}_{ij} = (\text{SUBSIDY}_{ij} / \text{CAPRATE}_{ij}) / \text{AVGHV}_{ij}$$

where all variables are as defined above.

Figure 5 plots the mean CAPSUB for each metropolitan area, from lowest to highest value. In the typical metropolitan area, the capitalized value of the subsidy for a house in a typical census tract is around 20 percent of house value. However, the standard deviation is about 6 percentage points and the interquartile range for the metropolitan area means is from 18-25 percent. There are 20 metropolitan areas in which the mean CAPSUB value is over 25 percent. Included in this group are the MSAs or CMSAs centered around Boston, New York City, Washington, D.C., Los Angeles, and San Francisco, so the list contains some of the largest and most densely populated areas in the country. Values for individual metropolitan areas are reported in Appendix B.

Figure 5 suggests that if the subsidy were fully capitalized and the tax benefits suddenly taken away, house prices in the typical area would fall by around 20 percent, with some places experiencing much larger declines and a few areas being less affected. While this represents a decline in demand due to the substitution effect associated with the change in relative price of housing, there will be a countervailing income effect associated with the lump sum redistribution of the \$1,815 per household savings from eliminating the subsidy. The lump sum transfer makes it clear that renters will be better off if the tax subsidy to home ownership is eliminated. The

$\text{COSTS} = (1 - \tau)\text{SUBSIDY} + M + \delta - \Pi^H$. The variable SUBSIDY is computed for each tract as described in equation (4). We use an average tax rate defined as subsidy per owner in the tract divided by the mean housing deductions in the tract to proxy for τ . Values for maintenance, economic depreciation, and housing price inflation are taken from Poterba (1992). These three measures are the same across all tracts, so that the variance in capitalization rates across metropolitan areas is due solely to the spatial variance in SUBSIDY values. The mean capitalization rate across metropolitan areas is 9.7 percent with a fairly tight interquartile range of 9.4-10.0 percent. Ten percent of the areas have cap rates below 9.1 percent, while another ten percent have cap rates above 10.2 percent. Our examination found the variance to be sensible with the larger, more densely populated areas having the lower cap rates and the smaller and more rural areas tending to have the highest cap rates.

same will be the case for some owners, especially those who do not experience much of a reduction in house price in exchange for the stream of annual payments of \$1,815. Whether the trade-off is worth it for these owners depends upon other factors such as their expected investment horizons and their discount rates. For owners in relatively high CAPSUB areas, the trade-off very likely will not be worth it.

Because our measure incorporates no general equilibrium effects of a change in policy, precise estimates of the ultimate impacts are not possible. However, our back-of-the-envelope measure strongly suggests that the wealth effects of eliminating the tax subsidy to owners, if fully capitalized, would be economically meaningful for the average area in America. Research by Tracy et al (1999) shows that, even with the spread of stock ownership, housing wealth still constitutes virtually all of household wealth for the typical owner household. And, for those owners with leverage above 80 percent, such as younger households or first-time homeowners, elimination of the subsidy could wipe out all home equity if the subsidy is capitalized, raising the specter of increased mortgage defaults [Kau et. al. (1992); Deng, Quigley, and Van Order (1996)].

Our second measure, of the proportion of income saved due to the tax subsidy, is simply

$$\text{INCSUB}_{ij} = \text{SUBSIDY}_{ij} / \text{AVGINC}_{ij}.$$

In this case, SUBSIDY and AVGINC values are available for each tract in each metropolitan area. Absent any capitalization of the subsidy or changes in interest rates, elimination of the tax benefits must lead to an increase in user costs. Mean INCSUB values by metropolitan area are plotted in Figure 6. In a typical area, the annual costs of owning would increase by about four percent of annual income. While three quarters of MSAs have mean INCSUB values of between

3-5 percent, in 10 MSAs the fraction would be more than 10 percent.²¹

In this case, a cash lump sum transfer is presumed to be received each year to offset the higher cash user costs of ownership. Creating a 'net INCSUB' measure in which the numerator of equation (8) is changed to equal 'SUBSIDY_{ij} - \$1,815' finds that owners in the typical metropolitan area are better off from eliminating the tax subsidy if there is no capitalization and there is lump sum redistribution of the savings. The median metropolitan area has a 'net INCSUB' value of -0.012, indicating that the lump sum transfer exceeds the increase in user costs by just over one percent of income each year. In one quarter of metropolitan areas, the lump sum transfer exceeds the rise in user costs by at least 3.5 percent of annual income. However, even factoring in the lump sum savings, 'net INCSUB' is in excess of +0.027 for one quarter of the metropolitan areas. Owners in these areas clearly would be made worse off if the tax subsidy were eliminated. These results simply reflect the highly spatially skewed nature of the program's benefit distribution in which large subsidies flow to a relatively few areas.

IV. Do the Results Reflect Spatial Variation in Income or House Value Alone?

Our analysis shows that the tax subsidy to owner-occupied housing is considerably skewed spatially. However, it is not yet clear precisely what drives these findings. Given the progressivity of the tax code, one would expect the subsidy to be positively correlated with income, a factor that could go a long way toward accounting for much of what we have found so far. Hence, a natural question to ask is whether the spatial subsidy distribution merely reflects

²¹Those areas are Barnstable-Yarmouth (MA), Honolulu, Los Angeles-Riverside-Orange County, New York-Northern New Jersey, Non-MSA Hawaii, Non-MSA Massachusetts, Salinas (CA), San Francisco-Oakland-San Jose, San Luis Obispo, and Santa Barbara-Santa Maria-Lompoc.

the higher marginal tax rates associated with higher income.

The answer is no, as the subsidy distribution is materially more regressive than the tax code is progressive. For example, if we rank tracts by mean household income and concentrate on the tracts containing the top 10 percent of households by mean income in the tract, we find the following. The households in these tracts pay 28 percent of all taxes we estimate are paid nationally,²² yet the owners in these tracts receive 38 percent of the nationwide subsidy. Thus, the program appears more regressive than the tax code is progressive.²³

This suggests that house value, not just income (or marginal tax rate), plays an important role in accounting for the spatial variance in subsidy flows. Because these variables are so strongly correlated, it is difficult to establish the precise fraction of variance in subsidy across tracts that each accounts for. However, the subsidy distribution is not merely a reflection of the progressivity in marginal tax rates. High income, high tax bracket owners tend to reside in disproportionately valuable houses so that the tax code is interacting with housing prices to generate an extremely skewed tax subsidy distribution – both across the income scale and across

²²The estimate of taxes paid is an output from the NBER's TAXSIM program. The calculation is made at the tract level analogously to how subsidy estimates are made. See the discussion at the beginning of the previous section for the details.

²³The same conclusion is reached if one looks at the top quarter of households and so forth. That said, it should be mentioned that we underestimate the progressivity of the income tax code. In the *Statistics of Income* data for 1989, the 10 percent of tax filers with the highest adjusted gross incomes paid 55 percent of all taxes. While there are several possible reasons for the discrepancy, a major one almost certainly has to do with the fact that we average across households within each tract. Our highest mean income at the tract level is \$225,000. The *SOI* data show that households with incomes of \$250,000 or more paid 17.5% of all taxes in 1989. Effectively, by averaging across households within a tract, we reduce the observed skewness in income. In general, the more heterogeneity in incomes within tracts, the greater will be our underestimate of the progressivity of the tax code. A second likely contributing factor is that the Census does not collect some income items that would contribute to the skewness of income at the high end, such as capital gains. The census also top codes income.

We do not think the differences in progressivity matter much for our analysis because even if by using the tract as our unit of observation we reduce the apparent skewness in taxes paid and subsidies received, we think the *difference* between the two is robust. Consequently, our statement that the regressivity of the housing subsidy is greater than the progressivity of the tax code would still apply.

locations within the country.²⁴

V. Conclusion

The value of the tax subsidy to owners of homes is very large. In aggregate, the figure was nearly \$164 billion in 1989. It has long been known that these benefits are skewed towards owners with high incomes and high house prices. This paper produced the first detailed analysis of how this program plays out spatially, both across and within states, and across and within metropolitan areas. Our results showed that the fact that owner-occupied housing is tax-favored relative to other assets leads to substantial redistribution within and across states. The subsidy distribution was found to be highly skewed across metropolitan areas, too, with the bulk of them experiencing small to modest negative transfers on average, and a few areas receiving very large positive net transfers under the program.

How the subsidy is distributed within individual metropolitan areas varies widely, too. In many smaller MSAs, especially those in the interior of the country, program benefits tend to be distributed fairly evenly across owners. This is not the case in most larger, more populous areas, in which benefits tend to be skewed towards a relatively small fraction of owners. The spatial pattern of how these benefits are distributed within the metropolitan area also differs considerably across areas. Residential segregation by income – in which those with the highest

²⁴This can be shown more formally. For example, a plot of INCSUB against the fitted values from a regression of INCSUB on the log of mean house value in the tract shows that high house value owners receive relatively large subsidy flows under the current code even relative to income. Stated differently, those tracts with high average house values are predicted to have greater subsidies relative to income. Moreover, if INCSUB then is plotted against the fitted values from the regression of INCSUB on the log of mean household income in the tract, we find that higher income tracts (which are high tax bracket tracts) have higher INCSUB ratios. The analogous regressions and plots using CAPSUB indicate that high income tracts have high subsidies even relative to house value and that the perpetuity value of the subsidy is an increasing function of house value as house value increases.

incomes, marginal tax rates, and house values – tend to live close to one another is evident in virtually all metropolitan areas. However, benefits generally are much more evenly spatially distributed within some metropolitan areas (e.g., Madison and Terre Haute) than in others (e.g., Philadelphia).

The implications of our findings are important whether or not the subsidy is capitalized into land values. If there is substantial capitalization, eliminating the subsidy is likely to reduce house prices on the order of 20 percent on average in the United States. However, there will be substantial variation in this fraction across metropolitan areas. The impact is likely to be greatest – 25 percent or more – in our largest and most populous coastal areas. If there is little or no capitalization, user costs of owning must increase. Our calculations indicate that the increase will amount to between 3 and 5 percent of annual income for the typical owner in the bulk of metropolitan areas. However, the increase will average a much larger 10 percent of annual income for owners in twenty metropolitan areas, with the nation's biggest among them.

These results may help explain why the subsidy to owner-occupied housing persists even though most states, metropolitan areas, households, and even homeowners probably would be better off if it were eliminated and replaced with a lump-sum refundable credit. Those that are worse off due to the program do not lose much while those that benefit comprise the major metropolitan areas in the U.S. and gain a substantial amount.

Finally, the spatial distribution of this major subsidy program is not merely a reflection of a rising marginal tax rate structure. While it certainly is true that the tax subsidy increases with household income (and tax bracket), the program itself is more regressive than the tax code is progressive – and possibly by a fairly wide margin, although our averaging across owners in a

tract does result in an underestimate of the progressivity of the income tax code.

Selected References

- Bruce, Donald, and Douglas Holtz-Eakin. "Apocalypse Now? Fundamental Tax Reform and Residential Housing Values," *mimeo*, Syracuse University, November 1999.
- Capozza, Dennis, Richard Green and Patric Hendershott. "Taxes, Mortgage Borrowing, and Residential Land Prices." In Henry Aaron and William Gale, eds. *Economic Effects of Fundamental Tax Reform*. Washington, D.C.: Brookings, 1996, pp. 171-210.
- Deng, Yongheng, John Quigley, and Robert Van Order, "Mortgage Default and Low Downpayment Loans: The Cost of Public Subsidy", *Regional Science and Urban Economics*, Vol. 26, no. 3 (1996): 263-285.
- Feldstein, Martin, ed. *The Effects of Taxation on Capital Accumulation*, University of Chicago Press (1987).
- Gyourko, Joseph, Andrew Haughwout and Robert Inman. "Public Wealth and Local Land Values," Wharton Real Estate Center Working Paper, University of Pennsylvania, November 1997.
- Gyourko, Joseph, and Richard Voith. "The Relative Price of Housing and Urban Form: Density and Sorting Effects," *Working Paper*, The Wharton School, University of Pennsylvania, August 2000.
- Hendershott, Patric and Joel Slemrod. "Taxes and the User Cost of Capital for Owner-Occupied Housing," *AREUEA Journal*, Winter 1983, 10 (4), 375-393.
- Kau, J.B., D.C. Keenan, W.J. Muller and J.F. Epperson, "A Generalized Model of Fixed Rate Mortgages", *Journal of Money, Credit, and Banking*, Vol. 24, no. 3, (1992): 279-299.
- King, Mervyn. "An Econometric Model of Tenure Choice and Demand for Housing as a Joint Decision," *Journal of Public Economics*, October 1980, 14 (2), 137-159.
- Mills, Edwin. "Dividing UP the Investment Pie: Have We Overinvested in Housing?," *Business Review*, Federal Reserve Bank of Philadelphia, March-April 1987.
- Poterba, James. "Tax Subsidies to Owner-Occupied Housing: An Asset Market Approach," *Quarterly Journal of Economics*, November 1984, 99 (4), 729-752.
- ". "Taxation and Housing: Old Questions, New Answers," *American Economic Review*, May 1992, 82 (2), 237-242.

- Rosen, Harvey. "Housing Decisions and the U.S. Income Tax," *Journal of Public Economics*, June 1979, 36 (1), 87-109.
- Sinai, Todd. "Are Tax Reforms Capitalized into House Prices?" University of Pennsylvania: *Wharton Real Estate Center Working Paper*, 1998.
- Tracy, Joseph, Henry Schneider, and Sewin Chan. "Are Stocks Overtaking Real Estate in Household Portfolios?", *Current Issues in Economics and Finance*, Federal Reserve Bank of New York: Vol. 5, No. 5 (April 1999), 1-6.
- U. S. Bureau of the Census. *Summary Tape File 3 (STF3)*, census tract-level data. Washington, D.C.: 1990.
- U. S. Department of Health and Human Services. *1998 Green Book Overview of Entitlement Programs*, Ch. 7, Aid to Families with Dependent Children and Temporary Assistance for Needy Families, Table 7.2 (see URL www.aspe.hhs.gov/98gb).
- U.S. Department of the Treasury. *Statistics of Income, 1989 Public Use Microsample*. Washington, D.C.: 1990.

Appendix A: Correcting for Self-Selection in Determining Itemization Status

The SOI data set we use to help deal with self-selection problems contains information from the tax returns of a 96,000 observation, income-stratified sample of all tax filers in 1989. With ideal data, we could assign the mean deduction for each category of taxpayer to the census data. However, administrative data such as that in the SOI only reports deductions if the family actually itemized. If the family had some potentially itemizable deductions, but chose to take the standard deduction instead, we observe none of their potential deductions. Since only families with a lot of deductions itemize, it is unlikely that the mean of the observed deductions, conditional on having any, is the same as the unconditional mean (which is the value we wish to impute).

There are two reasons why we might not observe any deductions in a particular category. The first, which has already been mentioned, is that deductions are observed only if the taxpayer has enough total deductions to make itemization worthwhile. The second is that a taxpayer may not want or have deductions in any one category even if his total deductions across categories are sufficient to make itemization worthwhile. Each factor generates the need for a sample selection correction of its own.

Thus, we would like to estimate the following,

$$(A.1) D_{ij} = \alpha_0 + \alpha_1 * f(AGI_i) + \alpha_2 * FSTAT_i + \alpha_3 * CHILD_i + \epsilon_{ij}$$

where D_{ij} are the deductions for family i in category j , $f(AGI_i)$ is a nonlinear function of adjusted gross income (AGI),²⁵ $FSTAT$ represents the family's tax filing status, and $CHILD$ corresponds to whether the family has children. However, we do not observe D_j . Instead, we observe D^*_j

²⁵We use a set of indicator variables for ranges of adjusted gross income.

which equals D_j under the following conditions:

$$D_j^* = D_j \text{ if } \Sigma D_j > \text{standard deduction and } D_j > 0$$

$$(A.2) \quad D_j^* = 0 \text{ if } \Sigma D_j > \text{standard deduction and } D_j \neq 0$$

$$D_j^* \text{ is not observed otherwise (} \Sigma D_j \leq \text{standard deduction).}$$

Our imputation problem arises for two reasons. First, $E[D_j^*] \neq E[D_j]$. Second, the sample of families for whom D_j is observed may have a different responsiveness of itemization to AGI, so that we would not be able to extrapolate our imputation out of the sample of itemizers. We, therefore, adopt the following strategy. The first step involves using the SOI data to estimate whether a family itemizes. For the entire U.S., this is done via a probit for the following model,

$$(A.3) \quad I_k = \beta_0 + \beta_1 * f(AGI_k) + \beta_2 * FSTAT_k + \beta_3 * CHILD_k + \beta_4 * STATE_k + \eta_k,$$

where $I_k = 1$ if family k itemizes and zero otherwise. The variable $STATE$ is an indicator for the state of residence. It is included because state residence affects the probability of itemizing due to differences in state tax rates, but it should not independently affect the likelihood of having deductions in other categories or the amount of those deductions.

The second step in the imputation procedure requires constructing the inverse Mills ratio for each family k . For future reference, this is labeled as

$$(A.4) \quad \hat{\lambda}(I_k) = \varphi(X_k \hat{\beta}) / \Phi(X_k \hat{\beta}),$$

where X_k is the vector of right-hand side variables from equation (A.3). The next step involves using a probit with the SOI data to estimate equation (A.5) for the entire U.S.,

$$(A.5) \quad C_{jk} = \gamma_0 + \gamma_1 * f(AGI_k) + \gamma_2 * FSTAT_k + \gamma_3 * CHILD_k + \gamma_4 * \hat{\lambda}(I_k) + \zeta_k,$$

where $C_{jk}=1$ if family k has positive deductions in category j and $\hat{\lambda}(I_k)$ is included to correct for

sample selection in the pool of itemizers.

From equation (A.5), we can then construct $\hat{\lambda}(C_{jk}) = \varphi(Z_{jk} \hat{\gamma}) / \Phi(Z_{jk} \hat{\gamma})$, where Z is the vector of right-hand side variables in (A.5). We estimate equation (A.6) via OLS,

$$(A.6) D_{jk} = \delta_0 + \delta_1 * f(AGI_k) + \delta_2 * FSTAT_k + \delta_3 * CHILD_k + \delta_4 * \hat{\lambda}(I_k) + \delta_5 * \hat{\lambda}(C_{jk}) + \varepsilon_{jk}.$$

In this specification, $\hat{\lambda}(I)$ corrects for sample selection due to only observing itemizers and $\hat{\lambda}(C)$ corrects for the selection arising from the decision to itemize in a given category.²⁶

To impute to the census data, we apply the estimates from the SOI to the census data and construct λ_{Io} and λ_{Co} values for each census family i . Next, the value for C_{ji} , the probability of having a deduction in category j is imputed, where j corresponds to deductible medical expenses, charitable giving, deductible interest expenses, or other deductible expenses. We then pick a random number from the uniform distribution on the interval $[0,1]$ for each family. If that number is less than C_{ji} , we predict the amount D_{ji} , the amount of deductions in category j , and impute it to the family. If the random number is greater than C_{ji} , we impute zero as the amount of deductions.

²⁶It bears mentioning that the SOI only reports state of residence for families with AGI under \$200,000. However, more than 98 percent of families with AGI above that level itemize, whereas only 28% of families below that threshold do. Thus, we assume that there is no sample selection problem for families with AGI over \$200,000 and estimate the last three steps only, leaving $\hat{\lambda}(I_k)$ out of the regression for these families.