

Model Review: Household Tax Burden Model (Property Tax Estimation)

Date	June 9, 2020
Contact	Research and Fiscal Analysis Division (RFA) Analyst: Kris Bitney; krisb@dor.wa.gov ; (360) 534-1532 Manager: Valerie Torres; ValerieT@dor.wa.gov ; (360) 534-1521
Model Purpose	The Household Tax Burden Model simulates household tax burdens under current law or alternative tax policies. The model estimates the major taxes Washington imposes on households, including the property tax. The model generates estimates within income quantiles and geographic regions. The model also estimates tax burden as a share of household income so we can estimate the progressivity or regressivity of Washington’s tax structure for households. We measure tax burden as the total tax imposed on a household by state and local sources. This document describes property tax estimation in the Household Tax Burden Model.
Data Sources	The data used for property tax estimation in this model includes: IRS Individual Income Tax Data County Property Tax Rolls US Census Bureau - American Community Survey
Requirements Model Used to Fulfill	This part of the model will fulfill these requirements in ESHB 1109 (2019): ¹ Sec. 137(c)(vii)(A) With respect to the final report of findings and alternatives submitted by the Washington state tax structure study committee to the legislature under section 138, chapter 7, Laws of 2001 2nd sp. sess.: (I) Update the data and research that informed the recommendations and other analysis contained in the final report; (II) Estimate how much revenue all the revenue replacement alternatives recommended in the final report would have generated for the 2017-2019 fiscal biennium if the state had implemented the alternatives on January 1, 2003; (III) Estimate the tax rates necessary to implement all recommended revenue replacement alternatives in order to achieve the revenues generated during the

¹ Washington State Legislature (2019). *HB 1109: Making 2019-2021 biennium operating appropriations*. (<https://app.leg.wa.gov/billsummary?BillNumber=1109&Year=2019&Initiative=false>).

Model Review: Household Tax Burden Model

Requirements Model Used to Fulfill, *continued*

2017-2019 fiscal biennium as reported by the economic and revenue forecast council;

(IV) For estimates above, estimate the impact on taxpayers, including:

- Tax paid as a share of household income for various income levels, and
- Tax paid as a share of total business revenue for various business activities

(V) Estimate how much revenue all of the revenue replacement alternatives in the report would have generated for the 2017-2019 Biennium if the state had implemented the alternatives on January 1, 2003, excluding any recommendations implemented before the effective date of this act.

Sec. 137(c)(vii)(B):

(II) Estimate how much revenue would have been generated for the 2017-2019 biennium if the 1 percent revenue growth limit on regular property taxes was replaced with a limit based on population growth and inflation if the state had implemented this alternative on January 1, 2003.

Questions for Technical Advisory Group

We greatly appreciate any and all feedback about the model. Specific question include:

Our home ownership prediction method over-estimates home ownership because average household incomes are higher in the IRS Individual Income Tax Data than in the American Community Survey. How can we improve our estimates?

We use a regression model to estimate taxable property value for households missing that data, described in this document. How can we improve our approach to taxable property value estimation?

Questions from Technical Advisory Group

We will capture at our meeting and record here

Model Technical Description

Summary

The Household Tax Burden Model simulates the tax burdens of Washington households under alternative tax policies or current law. The model estimates the major taxes imposed on Washington households. Among the most important of those is the property tax, the focus of this document.

To estimate household property tax burdens, we associate households with real property by joining two datasets—one which is representative of Washington Households and one which contains record-level tax data on all real property in Washington. Because we cannot identify with certainty which households own the properties where they reside, we estimate the conditional probability that each household owns their residence and assign home ownership accordingly. Not all households are successfully merged with a property record. Using households that we successfully matched to properties, we train a regression model and use it to estimate property values for households that are not successfully matched to properties. We then calculate property tax burdens based on property values. We aggregate and report final estimates by income quantile and geographic region. This document describes these procedures and related challenges in detail.

Objectives

Objective 1: Estimate and report household tax burdens, measured as the total tax imposed on a household by state and local sources.

Objective 2: Estimate household tax burden as a share of household income, and report the degree of proportionality of Washington's tax structure for households.

Objective 3: Model alternative rates and types of taxation.

Table 1 in the appendix contains the output of the 2014 Household Tax Burden Model. It is our objective to replicate that table—with updated results—for each geographic region in which we estimate tax burdens and for each potential tax structure change with first incidence on households.

About the Data

IRS Individual Income Tax Data

IRS Individual Income Tax data includes data collected from federal individual tax returns for all Washington residents who filed federal tax returns for the 2017 tax year. Tax returns include form 1040 and supplemental forms and schedules.

County Property Tax Rolls

The Property Tax Rolls contain data on property valuations for Washington residents. Washington counties provide the data to the Department of Revenue. The data provided differs by county. This model uses data from 2016 property valuations, which represents tax payments due in 2017. This aligns with the IRS individual income tax data used in the model.

Model Review: Household Tax Burden Model

American Community Survey

The American Community Survey, published by the US Census Bureau, is a survey of American households covering a variety of topics. Those topics include key household, family, and economic characteristics. We use the American Community Survey “5-year estimates” dataset that includes Washington households surveyed between 2013 and 2017. Using frequency weights designed to make it statistically representative of Washington, the data can be used to obtain estimates for local geographic regions within the state.

Data Assumptions and Considerations

Representativeness of the IRS Individual Income Tax Data

We use the IRS Individual Income Tax Data to represent the Washington population. In doing so, we assume households filing federal income tax returns are representative of Washington as a whole.

It is worth noting that estimates may underrepresent low-income and no-income households who are not required to pay federal income tax or choose not to file. The underrepresentation of low-income households may adversely affect the accuracy of our results, particularly our estimates of the taxes borne by the bottom quantile of income earners.

The unit of observation in the IRS Individual Income Tax Data is a tax return. A tax return does not necessarily correspond to a household as defined by other sources used in the model, such as the American Community Survey. To account for this, we collapse the IRS Individual Income Tax Data dataset on home address. In other words, we assume that everyone sharing an address belongs to the same household unit. Because some taxpayers live in institutional housing—e.g., colleges, military bases—we do not apply this procedure to tax returns listing P.O. Box addresses (~6.8% of returns) or when more than six tax returns are associated with a single address (a marginal increase of 0.4% of returns).

Variables in the American Community Survey

Table 2 lists the variables in common between the IRS Individual Income Tax data and the American Community Survey. Both the IRS Individual Income Tax Data and the American Community Survey include some income types that are not common to both datasets.

Table 2: Variables in both IRS Individual Income Tax Data and American Community Survey data

Variable Description
Family size
Presence of person under age 18 in household
Presence of person over age 64 in household
Geographic region within Washington (19 regions)
Household Income (sum of shared categories)

When relating the datasets using a statistical model, we are limited to the use of income categories the datasets have in common, individually or in aggregate. This is a limitation because, for example, our statistical model does not account for capital gains as a source of income. This may affect the accuracy of our estimates for some high income households. The limitation does not affect our reporting. We can report results based on the full income information available in the IRS Individual Income Tax Data.

Relationship between IRS Individual Income Tax Data and County Property Tax Rolls

To estimate property tax burdens, we join IRS Individual Income Tax Data and County Property Tax Rolls datasets using physical and mailing addresses. There is a time difference between the two datasets. Tax authorities assessed the property values in calendar year 2016, but we have tax filers’ addresses as of spring 2018, when they filed their TY2017 tax returns. We assume housing mobility between 2016 and 2018 will not meaningfully affect our results.

Methodology

Geocoding and address matching

We geocode and standardize addresses for households and parcels in the Individual Income Tax data and County Property Tax Rolls data. We then join households to parcels using address as a composite key. We limit the County Property Tax Rolls data to single-family units to increase the likelihood that households own the parcels to which they are matched.

We find parcel matches for about 56.2% of the 2,582,642 unique addresses in the IRS Individual Income Tax Data. Based on data from the 2017 American Community Survey, we estimate that about 67.4% of Washington households live in single-family units. This suggests we successfully matched about 83.3% of Washington households living in single-family units to parcels.

Of those households likely living in single-family units, about 16.7% do not match with parcels from the County Property Tax Rolls data—about 11.2% of all households in the IRS Individual Income Tax data. About 6.8% of tax returns list PO Boxes as their addresses, and we do not expect those addresses to match to parcels. If PO Box owners are as likely as other households to live in single-family units, then the presence of PO Box addresses in 6.8% of tax returns explains why about 4.3% of all households did not match to parcels. It is possible the remaining 6.9% do not match with parcels due to imperfections in the data or our process, but we cannot be sure of the cause.

Many additional households do not match to parcels because they live in multi-family units, such as apartment buildings, that we do not include in the County Property Tax Rolls data we use. Estimates from the American Community Survey suggest about 32.6% of Washington households lived in multi-family units in 2017.

We can also measure address matching performance as the share of taxable value for which we can account. We successfully matched about 66% of the 1,510,862 single-family unit parcels in the County Property Tax Rolls data to households in the IRS Individual Income Tax data, accounting for about 72% of total taxable value.

Property Tax Estimation

We generally calculate the property tax burden for each household using actual property value data. However, assigning property tax burdens to households can be challenging for two reasons. First, we do not know with certainty which households own their homes. Second, many households in the IRS Individual Income Tax Data are not successfully matched to County Property Tax Rolls parcels. We address these issues by using statistical models to predict home ownership and estimate property values when this information is not otherwise available.

Home ownership

Knowledge of home ownership is important because home ownership determines property tax liability. We have two sources of information that indicate home ownership for subsets of households.

From tax return data, we know which households paid mortgage interest during the 2017 calendar year. IRS Individual Income Tax forms, such as form 1098, do not tell us whether mortgage interest payments are associated with a primary residence or another property. We assume taxpayers are homeowners if they paid more mortgage interest than they claimed as rental real estate losses on form Schedule E.

Model Review: Household Tax Burden Model

An additional indicator of home ownership, found in the County Property Tax Rolls dataset, is Washington's property tax exemption program for senior citizens and disabled persons. Program eligibility requires ownership and occupation of a primary residence in Washington.

We use the American Community Survey to estimate the probability of home ownership for the remaining subset of households. We model the conditional probability of home ownership as a function of household income, family size, presence of a minor in the household, presence of someone over age 64 in the household, a proxy for living in a single family unit, and geographic region within Washington. We implement the conditional probability function using a probit classification model. The model is depicted in Equation 1, where $P()$ denotes probability, $F()$ is the normal cumulative distribution function, y is a binary indicator of home ownership, X is a matrix of covariates, and β is the vector of parameters to be estimated.

$$\text{Equation 1: } P(y|X) = F(X\beta)$$

After training the model using data from the American Community Survey, we use it to predict home ownership in the IRS Individual Income Tax data.

We can check the validity of this model by comparing predictions of home ownership to our knowledge of home ownership from other sources. Among households paying mortgage interest or participating in the property tax exemption program for senior citizens and disabled persons, the model predicts a home ownership rate of about 97%. The true home ownership rate among that population is assumed to be 100%, suggesting the model is reasonably accurate for this group.

We can also compare predicted home ownership rates in the IRS Individual Income Tax data to home ownership rates in the American Community Survey. The model overestimates the home ownership rate when applied to the full population of Washington households in the IRS Individual Income Tax data. Estimates from the American Community Survey suggest about 57.6% of Washington households owned their homes in 2017. We assume a household owns their home if their predicted probability of home ownership is at least 80%, leading to an estimated home ownership rate of about 57.5% in the IRS Individual Income Tax data. Were we to assume home ownership when the predicted probability of ownership is at least 50%, our overall estimated home ownership rate would be about 71%. We can largely attribute that overestimation of home ownership to higher average household incomes in the IRS Individual Income Tax data. Requiring greater model confidence leads to a more accurate home ownership rate.

Households not matched to parcels

Importantly, we do not have property tax information for households that do not match with parcels. The aforementioned home ownership prediction model suggests this data is missing for about 10% of all households that own their homes. We estimate property values for these households.

Using the sample of households we can successfully match to parcels, we train a linear regression model that estimates taxable property values for the households that do not successfully match to parcels. We model taxable property value as a function of household income, family size, senior exemption status, and county of residence. We include a natural cubic spline basis for household income, with knots placed where household income equals 5%, 27.5%, 50%, 77.5%, 95%, and 99% of the maximum value of household income in the sample. The model has an adjusted R^2 of 0.27. The model is depicted in Equation 2, where y is taxable property value, X is a matrix of covariates, and β is the vector of parameters to be estimated.

$$\text{Equation 2: } y = X\beta$$

Model Review: Household Tax Burden Model

After we estimate property values for the households missing them, we use the estimated property values to calculate tax burdens as if the true values were known.

Tax rates

We use actual levy and taxable property value data to estimate property tax rates. We estimate state and local property tax rates by county and region. We use the following levies in our calculations:

- State property tax levy, part 1
- State property tax levy, part 2
- Local regular property tax levy
- Local special property tax levy
- Local school property tax levy

Recent changes to Washington's state property tax system are worth noting.

The first part of the state property tax levy was calculated using a "budget-based" system prior to assessment year 2017. Tax districts could not increase regular property tax levies by more than 1% of the regular levy they imposed in the prior year, plus the tax rate applied to new construction and properties assessed by the state. In 2017, Washington adopted a "rate-based" property tax system. For taxes due between calendar year 2018 and 2021 (assessment years 2017-2020), the property tax rate in part 1 was held fixed.

The second part of the state property tax levy was first implemented for property taxes due in calendar year 2018 (assessment year 2017). It is calculated as the difference between part 1 of the state property tax levy and a target rate. For property taxes due in 2018, the target rate was \$2.70 per \$1,000 of property value. For taxes due in 2019, the target rate is \$2.40. The target rate returned to \$2.70 for taxes due in 2020 and 2021. The second part of the state property tax levy ensures the typical household property tax burden matches the target rate.

Washington is returning to a "budget-based" system in assessment year 2021, for taxes due in 2022. The second part of the levy will no longer be determined by a target rate. Increases in both levies will be limited to 1% of the levies imposed in the prior year, plus consideration for new construction and properties assessed by the state. Although the combined levies will be capped at \$3.60 per \$1,000 of property value, property tax rates tend to decline annually in Washington's budget-based system—as property values increase faster than the 1% levy increase limit—and the cap is unlikely to become relevant in the absence of new legislation.

We calculate and apply state and local property tax rates that were due in calendar year 2017 (assessment year 2016). After calculating household property tax burdens and aggregating by income quantile and region, we adjust total aggregated tax burdens to match known property tax totals to account for changes in tax rates and property values.

Note that the estimates for Sec. 137(c)(vii)(B), which refers to a change in the property tax revenue growth limit, will be statewide estimates only.

Expected Outputs

The primary model output is a table for each of the geographic regions for which we compute estimates. We use each table to describe the mean household tax burden for each major consumer tax, as well as the mean total household tax burden, by income quantile. The tables also describe tax burden as a share of income within each income quantile. We measure tax burden as the total tax imposed on a household from state sources.

Table A2 in the appendix contains the output of the 2014 Household Tax Burden Model. It is our objective to replicate that table—with updated results—for each geographic region and statewide in which we estimate tax burdens.

Resource Links

Data Sources

[American Community Survey Public Use Microdata Technical Documentation](#)

[IRS Prior Year Forms and Instructions](#)

[Washington State Economic and Revenue Forecast Council](#)

Table A1: Output of the 2014 Household Tax Burden Model

Tax Burden on Households Major State and Local Taxes

Current Law

Household Income	\$0	\$15,000	\$25,000	\$35,000	\$45,000	\$55,000	\$70,000	\$85,000	\$105,000	over
	\$15,000	\$25,000	\$35,000	\$45,000	\$55,000	\$70,000	\$85,000	\$105,000	\$140,000	\$140,000+
Retail Sales Tax	\$905	\$1,170	\$1,453	\$1,690	\$1,988	\$2,340	\$2,729	\$3,217	\$3,832	\$5,908
Alcoholic Beverages Taxes	\$55	\$63	\$78	\$93	\$98	\$119	\$125	\$143	\$167	\$239
Cigarette & Tobacco Taxes	\$156	\$177	\$194	\$197	\$204	\$211	\$198	\$193	\$166	\$119
Insurance Premiums Tax	\$22	\$36	\$47	\$55	\$67	\$74	\$83	\$94	\$107	\$138
Gasoline Tax	\$132	\$184	\$233	\$273	\$311	\$349	\$392	\$427	\$467	\$498
Public Utility Taxes	\$113	\$140	\$157	\$171	\$186	\$199	\$217	\$236	\$258	\$324
Property Tax	\$804	\$1,089	\$1,307	\$1,524	\$1,896	\$2,230	\$2,678	\$3,102	\$3,824	\$6,130
Total Tax	\$2,187	\$2,859	\$3,469	\$4,003	\$4,749	\$5,522	\$6,422	\$7,413	\$8,821	\$13,354
Tax as % of Income	26.5%	14.2%	11.5%	10.0%	9.5%	8.9%	8.3%	7.8%	7.3%	5.9%