Online Appendix For
"State Taxes, Migration and Capital Gains Realizations"

August 3, 2022

A Detailed Summary Statistics

Table A.1 provides basic summary statistics on the individuals in our primary sample.

Table A.1: Summary Statistics on Residents of High Tax States with Large Capital Gains Realizations

Panel A: Means and medians

	Mean	Median
Income	488,818	161,453
Wages	246,686	37,143
Realization amount	2,332,290	749,207
Age	59.8	60.0
Observations	552,900	552,900

Panel B: Common industries

Industry	*NAICS code*	*Share of obs*
Real Estate and Rental and Leasing	53	0.158
Finance and Insurance	52	0.107
Professional, Scientific, and Technical Services	54	0.106
Manufacturing	31-33	0.082
Health Care and Social Assistance	62	0.054
Retail Trade	44-45	0.048

Panel C: Common origin states

State	*Share of obs*
California	0.449
New York	0.155
New Jersey	0.067
North Carolina	0.046
Ohio	0.037
Georgia	0.036

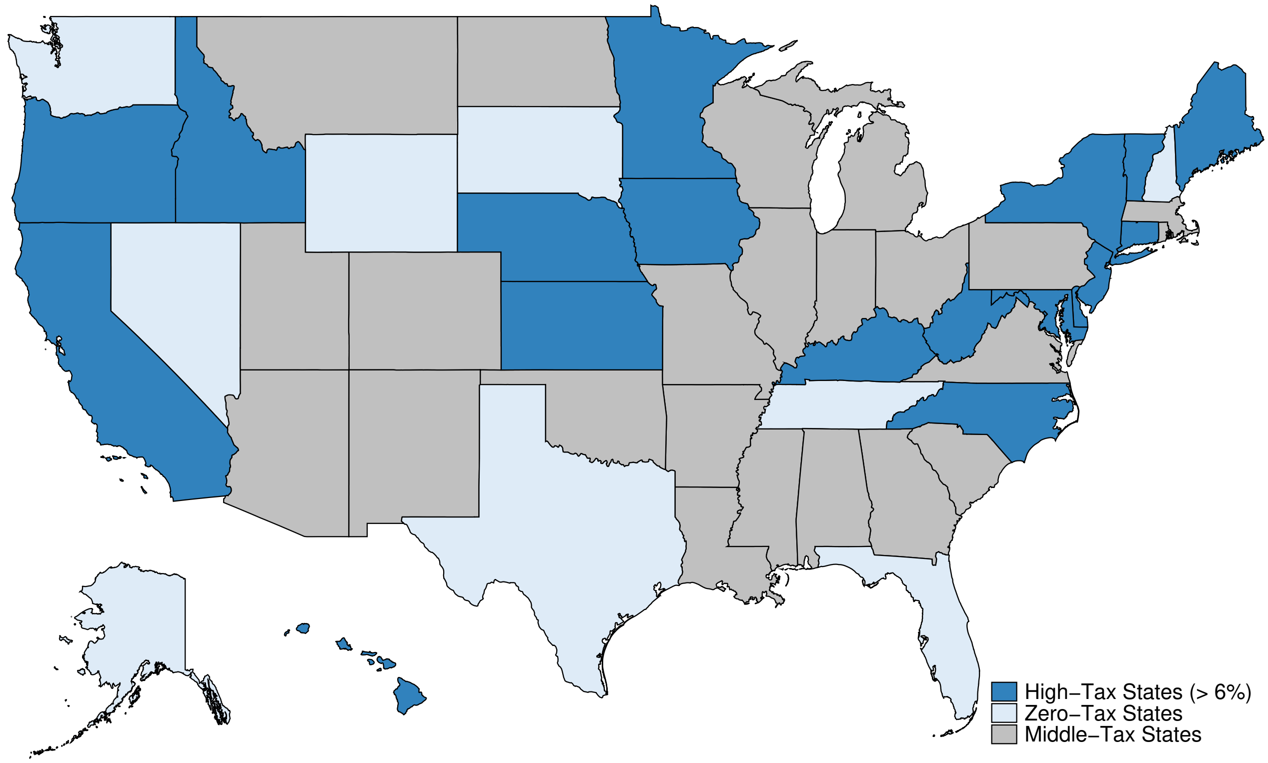
Panel D: Source of realization (2012)

Description	*Share of overall gain*
C corp stock	0.300
Passed-through gain	0.202
Pass-through interests	0.086
Other assets	0.082
Residential rental property	0.075
Depreciable real property	0.074
...	
Residences	0.004

Notes: This table provides summary statistics on our primary sample. Panel A provides information on realization sizes, realizer income and realizer age. Panel B reports the 2-digit NAICS code associated with the filer's primary source of income. (NAICS codes are collected in the three years prior to realization from Form W-2, and Schedule K-1 of Form 1065 and Form 1120S.) Panel C reports the most common origin states. Panel D provides information on the composition of capital gains amongst sub-sample of 2012 tax filers that were also analyzed in the Sales of Capital Assets study produced by the IRS Statistics of Income Program.

A map of US capital gains tax rates is displayed in Appendix Figure A.1 below:

Figure A.1: US State Capital Gains Tax Rates (2011)



Notes: This table shows top marginal tax rates in all US states in 2011. Data is from NBER TAXSIM (Feenberg and Coutts, 2018). States are grouped into three broad categories: 1) High-tax states – those with top tax rates above 6%, 2) Zero-tax states – those with no personal income taxes and 3) Middle-Tax States – all remaining states.

B Dynamic Discrete Choice Model Details

In this section we provide further detail on the dynamic discrete choice model described in Sections III and IV.

B.I Setting up the dynamic discrete choice model

We set up the dynamic discrete choice model as follows. In each period t, individual make two choices, they choose their state of residence and they choose whether they realize. Their choice set is:

$$C_{it} = (s_{it}, r_{it}) \tag{B.1}$$

- s_{it} captures the state that the individual chooses to live at in period t. We let $s_{it} = \{H_j, Z_j\}$, in other words, the individual can choose to be in a high tax state H_j or a zero tax state Z_j.
- r_{it} captures whether the individual realizes in period t. $r_{it} = \{0, 1\}$ is an indicator variable for realization. For tractability, we require that the individual realizes at some point in time, but place no restriction on when that realization occurs. We will then examine migration relative to that realization event.

There are a number of state variables in the model, given by the vector x_{it}. We let:

$$x_{it} = (z_{it}, s_{it-1}, \tau_{s_{t-1}} Q_i) \tag{B.2}$$

- z_{it} captures exogenous demographic characteristics regarding person i. This includes indicators for 10-year age bins, and 10 decile bins for income in advance of realization. Income is measured based on non-capital gains income (AGI - capital gains), 5-6 years prior to realization.

- s_{it-1} captures the individuals state of residence in the previous period. This is necessary to detect whether individuals have migrated.

- Q_i captures the size of the individuals' unrealized capital gains and $\tau_{s_{t-1}}$ captures the tax rate in state s, where the individual lived in period t-1. Together this state variable captures the taxes owed on unrealized capital gains. We use the lagged state of residence because individuals individuals who realize in one period pay capital gains based on their residence at the start of a given period.

Having established the choice set and the state variables, we set up the following flow payoff for each individual:

$$\pi_{it}(s, r) = z'_{it}(\alpha_{s,g} + \eta m_{it}(s_{it}, s_{it-1})) - \theta f(\tau_{s_{t-1}} Q_{it}) r_{it} \tag{B.3}$$

- $\alpha_{s,g}$ represents a vector of coefficients that can be estimated. The length of $\alpha_{s,g}$ is equal to the length of our demographic characteristics vector z_{it}. $\alpha_{s,g}$ captures the value of residing in the current state, $s$during tax regime g, for individuals with the various demographic characteristics captured in z_{it}. We classify tax regimes as periods of time where state tax rates do not vary by more than 1%. This is an attempt to capture variation in preferences for given locations that might vary with tax rates over time.

- $m_{it} = \mathbb{1}(s_{it} \neq s_{it-1})$ is an indicator that captures whether the individual has migrated since the previous period. η is the coefficient capturing the cost of migrating.

- $\tau_{s_{t-1}} Q_{it} r_{it}$ captures the state capital gains taxes paid. r_{it} is an indicator for whether realization occurred, $\tau_{s_{t-1}}$ captures the tax rate in state s (the individual's residence at the beginning of the period) and Q_{it} captures the size of the realization. $f(\cdot)$ captures the fact that the individuals payoff is a function of the tax savings. (We seek to estimate that function in Appendix B.IV below.) θ is our coefficient of interest.

Having established flow payoffs, we set up flow utility:

$$u_{it}(s, r, x) = \pi_{it}(s, r) + \varepsilon_{it}(s, r) \tag{B.4}$$

- Flow utility here is equal to flow payoffs plus an error term. We make the standard assumption that $\varepsilon_{it}(s, r)$ are i.i.d across i, s, r, and t with a Type I extreme value distribution.

Given that flow utility, the individual dynamic optimization problem is given by the following value function:

$$V(x_{it}) = \underset{s,r}{max}\{u_{it}(s, r, x_{it}) + \beta E[V(x_{it+1}(s_t, r_t))]\} \tag{B.5}$$

We also have a choice-specific (or conditional) value function given by:

$$\bar{V}(s_t, r_t, x_{it}) = \pi_{it}(s_t, r_t) + \beta E[V(x_{it+1}(s_t, r_t))] \tag{B.6}$$

This gives us the returns of an action before the realization of the error term, $\varepsilon_{it}(s, r)$.

B.II Implementing the Euler conditional choice probability approach

In order to use the Euler CCP approach in this context, we begin by drawing upon a version of Lemma 1 from (Arcidiacono and Miller, 2011):

$$V(x_{it}) = \bar{V}(s_t, r_t, x_{it}) - logPr(s_t, r_t|x_{it}) + \gamma$$

γ in this equation is Euler's Gamma. This equation is derived in the context of the Euler CCP in Scott (2013). It draws upon properties of the logit.

Once we have that lemma, we can use it to manipulate our equation for the choice-specific value function:

$$\begin{aligned}\bar{V}(s_t, r_t, x_{it}) &= \pi_{it}(s_t, r_t) + \beta E[V(x_{it+1}(s_t, r_t))] \\ &= \pi_{it}(s_t, r_t) + \beta E[\bar{V}(s_{t+1}, r_{t+1}, x_{it+1}(s_{t,} r_t)) - lnPr(s_{t+1}, r_{t+1}|x_{it+1}(s_{t,} r_t)) + \gamma]\end{aligned}$$

This equation holds for any set of choices (s_t, r_t) and any set of follow-up choices (s_{t+1}, r_{t+1}). In order to handle the continuation values that appear in this equation, we utilize the logic of renewal actions. (Scott (2013) notes that renewal actions are a special case of finite dependence.) For two sets of individuals who start in the same place and end in the same place, their value foward-looking functions can be equal even if they took different paths to get to that end point.

In our set-up we focus on two sets of individuals: those who migrate before they realize and those who realize before they migrate. For the individuals who migrate before they realize, the make the following choices: $(Z, 0)$ and then $(Z, 1)$. For the individuals who realize before they migrate, they make the following choices: $(H, 1)$ and then $(Z, 0)$. Both of these groups find themselves in the zero tax state after period 2 having realized their capital gains[1]. With that in mind we write out the choice specific value function for the initial choices along both of these paths:

$$\begin{aligned}&\bar{V}(s_t = Z_j, r_t = 0, x_{it}) = \pi_{it}(s_t = Z_j, r_t = 0) \\ &+ \beta E[\bar{V}(s_{t+1} = Z_j, r_{t+1} = 1, x_{it+1}(s_{t,} r_t)) - lnPr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0)) + \gamma]\end{aligned}$$

$$\begin{aligned}&\bar{V}(s_t = H_j, r_t = 1, x_{it}) = \pi_{it}(s_t = H_j, r_t = 1) \\ &+ \beta E[\bar{V}(s_{t+1} = Z_j, r_{t+1} = 0, x_{it+1}(s_{t,} r_t)) - lnPr(s_{t+1} = Z, r_{t+1} = 0|x_{it+1}(s_t = H_j, r_t = 1)) + \gamma]\end{aligned}$$

We can then take the difference between these two choice-specific value functions and utilize the fact that the continuation values: $V(x_{it+1})$ are nearly identical. In particular, $V(x_{it+1}) = EV(x_{it+1}) + \epsilon_t(s_t, r_t)$. The value function is equal to the expected value function plus an expectational error. And we can substitute in this equation, cancel out the equal value functions and only be left with an expectational error. We will call that expectational error $\epsilon_t(s_t = Z_j, r_t = 0) - \epsilon_t(s_t = H_j, r_t = 1) = \Delta\epsilon_t$. Putting it together, we get:

$$\begin{aligned}\bar{V}(s_t = Z_j, r_t = 0, x_{it}) - \bar{V}(s_t = H_j, r_t = 1, x_{it}) =& \pi_{it}(s_t = Z_j, r_t = 0) - \pi_{it}(s_t = H_j, r_t = 1) \\ &+ \beta[\pi_{it+1}(s_{t+1} = Z_j, r_{t+1} = 1) - \pi_{it+1}(s_{t+1} = Z_j, r_{t+1} = 0)] \\ &- \beta[ln\frac{Pr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0))}{Pr(s_{t+1} = Z_j, r_{t+1} = 0|x_{it+1}(s_t = Z_j, r_t = 1))}] \\ &+ \Delta\epsilon_t\end{aligned}$$

From there, we can utilize the following equivalence from Hotz and Miller (1993):

$$\bar{V}(s_t = Z_j, r_t = 0, x_{it}) - \bar{V}(s_t = H, r_t = 1, x_{it}) = ln[\frac{Pr(s_t = Z_j, r_t = 0|x_{it})}{Pr(s_t = H_j, r_t = 1|x_{it})}]$$

We can then combine that with the previous equation to get:

[1] The use of only two periods here is slightly stylized. In Section B.III to discuss how this maps onto our data.

$$
\begin{aligned}
&ln[\frac{Pr(s_t = Z_j, r_t = 0|x_{it})}{Pr(s_t = H_j, r_t = 1|x_{it})}] + \beta[ln\frac{Pr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0))}{Pr(s_{t+1} = Z_j, r_{t+1} = 0|x_{it+1}(s_t = H_j, r_t = 1))}] \\
&= [\pi_{it}(s_t = Z_j, r_t = 0) - \pi_{it}(s_t = H_j, r_t = 1)] - \beta[\pi_{it+1}(s_{t+1} = Z_j, r_{t+1} = 1) - \pi_{it+1}(s_{t+1} = Z_j, r_{t+1} = 0)] + \Delta\epsilon_t
\end{aligned}
$$

We can then use our previous formula for the flow payoffs to simplify the right hand side:

$$
\begin{aligned}
&[\pi_{it}(s_t = Z_j, r_t = 0) - \pi_{it}(s_t = H_j, r_t = 1)] - \beta[\pi_{it+1}(s_{t+1} = Z_j, r_{t+1} = 1) - \pi_{it+1}(s_{t+1} = Z_j, r_{t+1} = 0)] \\
&= z'_{it}(\alpha_{H_j,g} - \alpha_{Z_j,g}) + (z'_{it} - \beta z'_{it+1})\eta + \theta\left(f(\tau_H Q_{it}) - f(\tau_Z Q_{it})\right) + \Delta\epsilon_t
\end{aligned}
$$

Putting it all together we get the following equation:

$$
\begin{aligned}
&ln[\frac{Pr(s_t = Z_j, r_t = 0|x_{it})}{Pr(s_t = H_j, r_t = 1|x_{it})}] + \beta[ln\frac{Pr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0))}{Pr(s_{t+1} = Z_j, r_{t+1} = 0|x_{it+1}(s_t = H_j, r_t = 1))}] \\
&= z'_{it}(\alpha_{H_j,g} - \alpha_{Z_j,g}) + (z'_{it} - \beta z'_{it+1})\eta + \theta\left(f(\tau_H Q_{it}) - \beta f(\tau_Z Q_{it})\right) + \Delta\epsilon_t
\end{aligned}
\tag{B.7}
$$

B.III Working toward a linear regression to estimate the causal effect of tax savings

Next, we take the equation presented above and re-arrange in a way that simplifies it for direct estimation of the terms and the subsequent use of OLS.

In order to develop clearer intuition, we rearrange the left hand side of this equation (and set $\beta = 1$)[2]:

$$
\begin{aligned}
&lnPr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0)) + lnPr(s_t = Z_j, r_t = 0|x_{it}) \\
&- lnPr(s_{t+1} = Z_j, r_{t+1} = 0|x_{it+1}(s_t = H_j, r_t = 1) - lnPr(s_t = H_j, r_t = 1|x_{it})
\end{aligned}
$$

Which becomes:

$$
\begin{aligned}
&ln[Pr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0)) * Pr(s_t = Z_j, r_t = 0|x_{it})] \\
&- ln[Pr(s_{t+1} = Z_j, r_{t+1} = 0|x_{it+1}(s_t = H_j, r_t = 1) * Pr(s_t = H_j, r_t = 1|x_{it})]
\end{aligned}
$$

Here we can utilize the fact that $Pr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0))$ can be expressed as $Pr(A|B, C)$ and $Pr(s_t = Z_j, r_t = 0|x_{it})$ can be expressed as $Pr(B|C)$. Multiplying them together, we get $Pr(AB|C)$. Applied to this problem, we get: $Pr(s_t = Z_j, r_t = 0, s_{t+1} = Z_j, r_{t+1} = 1|x_{it})$. In other words, this is the probability of migrating in the first period and realizing in the second period, conditional on having the state variables in x_{it}. In other words, it is the probability of migrating in the first period and realizing in the second, conditional on having an initial residence in a high tax state and having a future realization of size Q.

So, we can re-write our left hand side as the following:

$$
ln[Pr(s_t = Z_j, r_t = 0, s_{t+1} = Z_j, r_{t+1} = 1|x_{it})] - ln[Pr(s_t = H_j, r_t = 1, s_{t+1} = Z_j, r_{t+1} = 0|x_{it})]
$$

[2] β must be calibrated in this dynamic discrete choice exercise. For the purposes of our calculations we assume that $\beta = 1$. The structure of our set-up means that small modifications to β will have a minimal impact on our results. In our primary regression the β will fall out of tax savings terms because, in Section B.IV, we impose the assumption that $f(0) = 0$. In other words, tax savings should not have an impact on an individual's behavior if they don't have any tax savings. If we were to impose the assumption that β is less than 1, we would need to alter our analysis such that individuals who migrate and then realize \$$Q$ are compared to individuals who realize slightly less than \$$Q$ before migrating. Our analysis in Section C.I, however, shows that this modification has very little impact on our results.

Or more simply:

$$ln\left[\frac{Pr(s_t = Z_j, r_t = 0, s_{t+1} = Z_j, r_{t+1} = 1|x_{it})}{Pr(s_t = H_j, r_t = 1, s_{t+1} = Z_j, r_{t+1} = 0|x_{it})}\right]$$

This gives us the final equation:

$$ln\left[\frac{Pr(s_t = Z_j, r_t = 0, s_{t+1} = Z_j, r_{t+1} = 1|x_{it})}{Pr(s_t = H_j, r_t = 1, s_{t+1} = Z_j, r_{t+1} = 0|x_{it})}\right] = z'_{it}(\alpha_{H_j,g} - \alpha_{Z_j,g}) + (z'_{it} - z'_{it+1})\eta + \theta\left(f(\tau_H Q_{it}) - f(0)\right) + \Delta\epsilon_t \tag{B.8}$$

On the left-hand side we have the log-odds ratio, which compare the probability of migrating before realization and the probability of realizing before migration. On the right hand side, $z'_{it}(\alpha_{H_j,g} - \alpha_{Z_j,g})$ captures individual preferences for residing in the high tax and low tax states; $(z'_{it} - z'_{it+1})\eta$ captures time-varying demographic differences between those taking each path, and $\theta\left(f\left(\tau_{H_j} Q_{it}\right) - f\left(0\right)\right)$ captures the role of potential tax savings.

B.IV Estimating a linear regression to estimate the causal effect of tax savings

In order to estimate the linear regression in Equation B.8[3], we need to take several key steps to map our model onto the data.

First, the regression in Equation B.8 is calculated conditional on the state variable x_{it}. In our primary specification we include a large number of state variables in x_{it}. We let $x_{it} = (z_{it}, s_{it-1}, \tau_{s_{t-1}} Q_i)$, we classify s_{it-1} as an indicator for whether an individual is in a high-tax state or a zero-tax state. We group $\tau_{s_{t-1}} Q_i$ into equal sized bins ranging from <\$50k for the lowest bin to \$1.5 million+ for the highest bin.[4] We measure capital gains realizations based on the size of an individual's largest realization. We use data created by NBER TAXSIM to calculate the top marginal tax rate in each state (Feenberg and Coutts, 2018). We let z_{it} represent individual specific variables and pre-period income. We group individuals into 10-year age bins. We also group individuals based on the size of their non-capital gains income (AGI - capital gains) in years 5 and 6 prior to realization.

Second, when it comes to migration we are interested in the choice to stay in a high-tax state or migrate to a zero-tax state. That said, we observe individuals in a number of high-tax states and a number of zero tax states. So in our primary specification we add richness to the model by focusing on the pairwise choices to stay in or migrate to individual states. In other words, we examine $Pr(s_t = Z_j, r_t = 0, s_{t+1} = Z_j, r_{t+1} = 1|x_{it})$ where the j indicates that we are calculating the probability of migrating to a specific zero-tax state such as Florida. In this case we calculate the left hand side probabilities for all pairwise combinations of origin and destinations. We then calculate our coefficients of interest on the pooled set of all realization and migration choices at various origins and destinations. This allow us to incorporate location-by-demographic preferences for each origin and destination location. We also look to account for differences in location preferences over time. In particular, we are concerned about changes in location preferences that are correlated with changes in tax rates. Those types of changes will impact our measure of tax savings and consequently might impact our estimate for θ. We solve this problem by examining migration choices within "tax regimes". We consider a tax regime to be any period of time in which a state does not change its tax rate by more than 1%. We allow location specific-preferences such as $\alpha_{H_j,g}$ to vary by both location and tax regime.

This version of our primary regression with a full set of controls lies in contrast to the simplest version of our regression equation where our state variables contain no demographic variables, z_{it}. Instead, individuals are merely grouped into tax savings bins, $\tau_{s_{t-1}} Q_i$ and individuals simply choose between residing in high or zero tax states. In that case, our primary regression equation is as follows:

$$ln\left[\frac{Pr(s_t = Z_j, r_t = 0, s_{t+1} = Z_j, r_{t+1} = 1|x_{it})}{Pr(s_t = H_j, r_t = 1, s_{t+1} = Z_j, r_{t+1} = 0|x_{it})}\right] = \theta\left(f(\tau_H Q_{it}) - f(0)\right) + \Delta\epsilon_t \tag{B.9}$$

[3] This is identical to Equation 4 in the body of the paper.

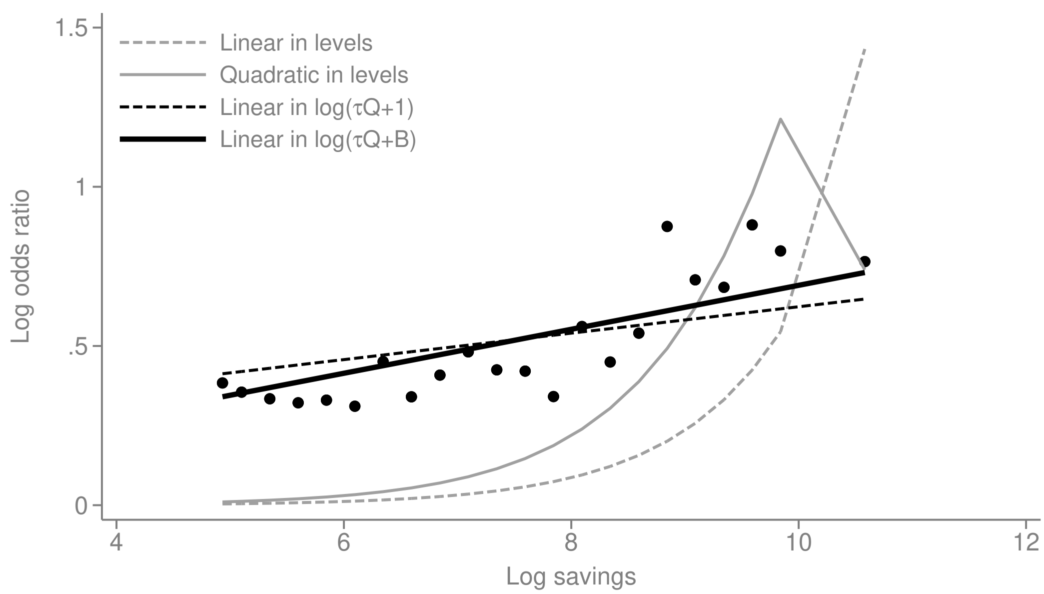
[4] We report tax savings in 2014 dollars.

As discussed in Section V in the body of the paper, this regression produces very similar results to our primary regression specification.

Third, we have focused our model two-period problem. Individuals realize and migrate in consecutive periods. In reality, we observe individuals migrating and realizing over the course of several years. We don't to throw out an observation simply because an individual migrated and then realized three years later. We address this problem by measuring time relative to migration or realization and then assuming that periods extend for four years. This avoids the need to incorporate complicated paths of renewal actions and also avoid complication associated with overlapping actions.[5] The choice of the four year periods is motivated by Figure I, which shows that the incentive to migrate before realization appears to occur within that time window. Limited foresight appears to attenuate the treatment effect in previous years.

Fourth, we are interested in calculating θ, our coefficient on $f\left(\tau_{H_j}Q_{it}\right) - f(0)$, which captures the impact of potential tax savings on individual payoffs. Thus far, we have written the expression $f\left(\tau_{H_j}Q_{it}\right) - f(0)$ rather than $\tau_{H_j}Q_{it} - \tau_{Z_j}Q_{it}$ in order to remain agnostic about the appropriate functional form in the context. Appendix Figure A.2 shows non-parametric relationship between tax savings and the log-odds ratio. This is clearly non-linear, and so the figure also shows the fit of a variety of different functional forms. Informed by this evidence, we estimate $f(X)$ as the following log function: $ln(X+B) - ln(B)$, where $B = 152$. As shown in Appendix Table A.2, this is the value that of B that minimizes the RMSE of the model. This specific functional form allows us to capture the log relationship in the data while also imposing the requirement that $f(0) = 0$. In other words, we impose the requirement that tax savings do not impact an individual's payoff function if they have no savings.

Figure A.2: Relationship Between Tax Savings and Migration Relative to Realization



Notes: This figure plots the relationship used to estimate the function, $f(\cdot)$ in Equation B.9. The Y-axis corresponds to the left-hand side of that regression, the log odds ratio comparing migration in advance of realization to realization in advance of migration. The X-axis displays potential tax savings. Individuals are grouped into potential savings bins and plotted based on the log-odds ratio within each bin. The figure then shows the parametric relationship between these two variables under a number of functional form assumptions. It provides a visual representation of the goodness of fit associated with $f(\cdot)$ being a linear function, a quadratic function, or a log function.

[5] If an individual migrates in year t and realizes in year t+2, we want to consider that realization that occurs after migration. If the timing of realization weren't measured relative to the timing of migration, it is possible these actions would be considered to take place in the same period.

Table A.2: Goodness of Fit for Values of B under the assumption that $f(X) = log(X + B)$

	(1) Log, B = 1	(2) Log, B = 50	(3) Log, B = 142	**(4) Log, B = 152**	(5) Log, B = 162	(6) Quadratic in τQ
RMSE	0.070787	0.064333	0.063436	**0.063429**	0.063431	0.284069

Notes: This table shows the root mean squared error associated with various values of B when $f(X)$ in Equation B.9 is set equal to $log(X + B)$. The goal of these results is to demonstrate that we've chosen the value of B that minimizes the RMSE. This value of B is then used to define the function form of $f(X)$ in all subsequent analysis.

Finally, in order to estimate this regression, we must calculate the ratio of probabilities on the left hand side. The left hand side here is calculated conditional on x_{it}, and so it can produce relatively narrow bins if we use a high-dimensional set of state variables. As Almagro and Domínguez-Iino (2021) discuss, the limited number of observations in each cell can produce bias in our results. (Scott (2013) also discusses how probabilities equal to 0 or 1 create concerns when applying the Hotz-Miller inversion necessary to implement the Euler CCCP approach.) For that reason, we need to find a way to produce a smoothed estimate of this log odds ratio. We follow the approach of Kalouptsidi et al. (2020) and estimate the left hand side conditional choice probabilities using a logit model. We estimate a logit that captures the relative probability of migrating before realization relative to realizing before migrating. We add indicators for individual tax savings bins and then add origin and destination fixed effects interacted with income and age bins. This allows us to predict the log odds ratio within each tax savings bins, smoothing over incomes, ages, origin states and destination states.[6] As Table A.3 below shows, this appears to have a relatively limited effect on our results. Introducing the smoothed regression with a bevy of controls produces estimates that are similar to our estimates in the case where we use no controls and estimate the left-hand side conditional choice probabilities directly. We also use an alternate method for predicting conditional choice probabilities that is informed by Diamond et al. (2017). In that paper, they use a Gaussian kernel to smooth over variables such as distance and age. In our case, we predict the probabilities of migrating in advance of realization using Gaussian processes. We take the prediction produced by the Gaussian process classification and use that in our OLS regression. As shown in Table A.3, this produces qualitatively similar results to our primary specification but the treatment effect is reduced by about 50%.

[6]We use origin and destination fixed effects instead of origin by destination fixed effects. For any given origin this allows us to smooth over the probabilities across destination and vice versa.

Table A.3: Impact of Conditional Choice Probability Smoothing

	No controls (1)	**Limited controls** (2)	**Baseline** (3)	**Gaussian Kernels** (4)
θ Coefficient	0.0627	0.0926	0.0965	0.0465
	[0.0593, 0.0664]	[0.0703, 0.1146]	[0.0750, 0.1194]	
Zero-Tax Realizations	\$2,021m	\$2,682m	\$2,846m	\$1,635m
	[1869, 2174]	[2166, 3143]	[2339, 3332]	
CA Fiscal Externality	\$14.3m	\$21.0m	\$13.7m	\$6.6m
	[13.1, 15.5]	[15.7, 26.5]	[10.5, 17.3]	
Income/age controls		X	X	X
Origin controls			X	X
Destination controls			X	X
Smoothed dependent variable		X	X	X

Notes: This table shows how the primary estimates in the paper are impacted by the use of alternate methods to calculate the conditional choice probabilities in the primary regression equation. Column (1) shows the results associated with estimating Equation B.9. In this case, the conditional choice probabilities are estimated directly in the data. Column (2) shows the results associated with estimating Equation B.8. In this case, additional age and income controls have been added to the regression and the conditional choice probabilities are predicted using a logit. This is the approach used in Kalouptsidi et al. (2020). Column (3) adds origin and destination specific controls to account for differences in origin-specific and destination-specific location preferences. Column (4) shows the results associated with estimating Equation B.8 and smoothing the conditional choice probabilities using Gaussian Kernels to smooth over age and income. This is the approach taken in Diamond et al. (2018). All 95% confidence intervals are obtained using the Bayesian bootstrap.

B.V Impact of Zero-Tax Opportunities on Realizations

In Section IV.B of the paper we examine a counter-factual where residents of high-tax states are not permitted to avoid capital gains taxation by migrating to zero tax states. In the body of the paper we refer to individuals as taking action A if they move and then realize. We refer to all other courses of action as part of a set of all paths J. For ease of explanation we now break the components of J. In particular, we split individual behavior into three potential actions.

- Individuals who take Action A move and then realize: $C_{it} = (Z_j, 0); C_{it+1} = (Z_j, 1)$.
- Individuals who take Action B realize and then move: $C_{it} = (H_j, 1); C_{it+1} = (Z_j, 0)$.
- Individuals who take any other course of action are referred to as taking Action C. For initial tractability we restrict this to the set of actions where individuals who begin in high-tax state realize their gains within 2 periods. We discuss these various courses of action in more detail below.

We are interested in identifying how the probability of Action A changes with changes in the tax rate, $P(A|x, \tau_Z = \tau_2) - P(A|x, \tau_Z = 0)$. The properties of the logit tell us that we can write $P(A|x, \tau_z = 0)$ in the following manner:

$$P(A|x, \tau_Z = 0) = \frac{exp(\bar{V}(A, x_i, 0))}{exp(\bar{V}(A, x_i, 0)) + exp(\bar{V}(B, x_i, 0)) + exp(\bar{V}(C, x_i, 0))}$$

Based on the payoff functions above, we can write out the choice-specific value functions for actions A and B :

$$\bar{V}(A, x_i, \tau_Z = 0) = z_i'\alpha_Z + z_i'\alpha_Z + z_i'\eta - \theta f(0) + E[V(x_{i,t+2})]$$
$$\bar{V}(B, x_i, \tau_Z = 0) = z_i'\alpha_H + z_i'\alpha_Z + z_i'\eta - \theta f(\tau_H Q_{it}) + E[V(x_{i,t+2})]$$

We can write the probability of taking action A in the following manner:

$$P(A|x,\tau_Z=0)=\frac{exp(\bar{V}(A,x_i,\tau_Z=0))}{\substack{exp(\bar{V}(A,x_i,\tau_Z=0))+\lambda_{BA}exp(\bar{V}(A,x_i,\tau_Z=0))\\ +exp(\bar{V}(C,x_i,\tau_Z=0))}} \tag{B.10}$$

Here, λ_{BA} is determined by the following expression: $exp(\bar{V}(B,x_i,\tau_Z=0))=\lambda_{BA}exp(\bar{V}(A,x_i,\tau_Z=0))$. In order to solve for λ_{BA} we can utilize the fact that if $exp(\bar{V}(B,x_i,\tau_Z=0))=\lambda_{BA}exp(\bar{V}(A,x_i,\tau_Z=0))$ then $\lambda_{BA}P(A|x,\tau_Z=0)=P(B|x,\tau_Z=0)$. We observe $P(B|x,\tau_Z=0)/P(A|x,\tau_Z=0)$.

In order to calculate the counterfactual probability $P(A|x,\tau_Z=\tau_2)$, we also need to know the relationship between $exp(\bar{V}(C,x_i,\tau_Z=0))$ and $exp(\bar{V}(A,x_i,\tau_Z=0))$. We write that relationship as $exp(\bar{V}(C,x_i,\tau_Z=0))=\lambda_{CA}exp(\bar{V}(A,x_i,\tau_Z=0))$. We can plug $\lambda_{CA}exp(\bar{V}(A,x_i,\tau_Z=0))$ into the denominator in equation B.10 above and solve for λ_{CA}. In our primary calculation we calculate $P(A|x,\tau_Z=0)$ as the fraction of individuals who move to a zero-tax state and then realize amongst the full set of individuals who begin in a high-tax state and realize within two periods.[7]

From there, we can consider the counterfactual probability $P(A|x,\tau_2)$, using updated choice-specific value functions. Following that approach, we get:

$$P(A|x,\tau_Z=\tau_2)=\frac{exp(\bar{V}(A,x_i,\tau_Z=\tau_2))}{exp(\bar{V}(A,x_i,\tau_Z=\tau_2))+exp(\bar{V}(B,x_i,\tau_Z=\tau_2))\ +exp(\bar{V}(C,x_i,\tau_Z=\tau_2))}$$

$$P(A|x,\tau_Z=\tau_2)=\frac{exp(\bar{V}(A,x_i,\tau_Z=0))exp(-\theta\left(f(\tau_2 Q_{it})-f(0)\right))}{\substack{exp(\bar{V}(A,x_i,\tau_Z=0))exp(-\theta(f(\tau_2 Q_{it})-f(0)))+\lambda_{BA}exp(\bar{V}(A,x_i,\tau_Z=0))\\ +\lambda_{CA}exp(\bar{V}(A,x_i,\tau_Z=0))}}$$

We can rearrange this equation to cancel out all instances of $exp(\bar{V}(A,x_i,\tau_Z=0))$ and get a direct estimate for $P(A|x,\tau_Z=\tau_2)$. It is worth noting that there are a few key assumptions needed in order calculate these counter-factual probabilities. First, this calculation requires that the continuation value functions across actions A and B are not impacted by the change in tax rate in the high tax state. This allows us to re-write $exp(\bar{V}(A,x_i,\tau_Z=\tau_2))$ as a multiplicative function of $exp(\bar{V}(A,x_i,\tau_Z=0))$ as $E[V(x_{i,t+2})]$ does not change. We consider this assumption to be reasonable as individuals taking both actions have found themselves in the zero tax state and have both realized their gains by the time they arrive at period $t+2$. The only difference between these cases are that the individuals have slightly different levels of wealth, as determined by their tax savings. Second, this calculation requires the assumption that $exp(\bar{V}(C,x_i,\tau_Z=0))=exp(\bar{V}(C,x_i,\tau_Z=\tau_2))$. In other words, changes in the tax rate don't change the value function for individuals who take a course of action other than Action A and Action B. This is intuitive in cases where individuals realize in the high tax state. Their value functions are not impacted by the policy change regardless of whether they stay in their home state after realizing or migrate to a non-zero tax state after realizing. The situation gets a bit more complicated when considering the value function for individuals who migrate to a non-zero but lower tax state and realize in that state. The current counter-factual assumes migrants to non-zero tax states can no longer avoid capital gains tax. We discuss the plausibility of this assumption in Section C.Vand explain why it has a limited impact on our results.

In order to translate these estimates into a quantity of realizations, we need to sum over individuals with a range of different demographic characteristics and realization quantities. For each set of $x's$ we calculate the probability of moving and then realizing. We use this to calculate the increase in the number of individuals who move and then realize. From there, we multiply that figure by the size of the individual's initial realization quantity. We estimate the total quantity of new realizations using the following expression:

$$\sum_Q\sum_x\left(\underbrace{N_{Q,x}Q}\underbrace{(P(A|x,\tau_Z=\tau_H)-P(A|x,\tau_Z=0))}\right) \tag{B.11}$$

This is the calculation we conduct to produce the \$2.8 billion estimate reported in Section IV.A. It is also worth noting that counterfactual we evaluate here is similar, but not identical, to a policy equalizing

[7]In the discussion that follows, we explore the implications of expanding that denominator. In particular, we consider there where some individuals who have state variables x_{it} but may substantially delay their realization. We show that adjustment has no substantive impact on our results.

capital gains tax rates across states. One key difference is that we don't assume that all residents of high-tax states are subject to the same tax rate. Residents of New Jersey continue to pay a top rate of nearly 9%, while residents of Oregon continue to pay a top rate of nearly 10%. We simply assume that individuals who accumulate unrealized gains in a high-tax location must pay taxes on those gains in their high-tax location. We believe that the counter-factual we analyze has more practical significance as changes in state tax rules could move policy in that direction. Total equalization of state taxes would produce a similar effect amongst highly mobile realizers, but such a policy change is infeasible.

As mentioned above, the calculation conducted here restricts our analysis to the set of individuals who begin in a high tax state and realize within the two periods. This is the sample that was necessary to calculate our coefficient of interest θ via the calculation of $P(A)/P(B)$. It is important to note, however, that some individuals with state variables x_t who are not observed realizing in this two period window. This includes, for example, individuals who maintain residence in their high-tax state for multiple periods before realizing. In order to confirm the robustness of our results we explore whether restricting our sample in this manner could impact our findings. In particular, we adjust our observed probabilities to account for the presence of these individuals. In order to do this, we need an estimate for the ratio of the number of individuals in our considered sample to the number of individuals with with state variable state variables x_t and realization quantity Q. We produce an estimate of this number by looking at individuals who reside in a high tax state in the first year of our data, 2002, and realize in 2003 or later. The incorporation of individuals who realize in any year between 2003 and 2018 allows us to capture the full set of individuals with the relevant state variables. For example, it allows for the possibility that some individuals who adjust away from Action A switch to a course of action where they hold their assets for 2 more periods before realizing in their home state.[8] Adjusting for the probabilities in this manner has little impact on our counter-factuals. We find that the zero-tax opportunities result in \$2.9 billion in additional realizations rather than the \$2.8 billion found in our primary specification.[9]

B.VI Reducing Top Tax Rates – CA Example

In Section IV.C of the paper we explore the impact of reducing top tax rates in California. In that case we evaluate a new counter-factual. We are again interested in the reduction in the number of individuals who migrate and then realize in zero tax states. In that case, we examine how behavior changes if top tax rates τ_{CA} falls by 1%.

Relative to Section 5.1 above we need to modify the set of actions taken by the individual:

- Action A occurs when individuals migrate to a zero tax state and then realize, their choice-specific value is $\bar{V}(A, x_i, \tau_H = \tau_{CA}) = z_i'\alpha_Z + z_i'\alpha_Z + z_i'\eta - \theta f(0) + E[V(x_{i,t+2})]$
- Action B occurs when individuals realize then migrate to a zero tax state, their choice-specific value is $\bar{V}(B, x_i, \tau_H = \tau_{CA}) = z_i'\alpha_H + z_i'\alpha_Z + z_i'\eta - \theta f(\tau_{CA} Q_{it}) + E[V(x_{i,t+2})]$
- Action C occurs when individuals remain in their home state and then realize, their choice-specific value function is $\bar{V}(C, x_i, \tau_H = \tau_{CA}) = z_i'\alpha_H + z_i'\alpha_H - \theta f(\tau_{CA} Q_{it}) + E[V(x_{i,t+2})]$
- Action D occurs when individuals realize in their home state and then move to a non-zero tax state, their choice-specific value function is $\bar{V}(D, x_i, \tau_H = \tau_{CA}) = z_i'\alpha_H + z_i'\alpha_M + z_i'\eta - \theta f(\tau_{CA} Q_{it}) + E[V(x_{i,t+2})]$
- Action E represents all other course of action. We discuss these courses of action in more detail below. Just as in Section B.V, we restrict our focus to courses of action where individuals realizing within 2 periods.

We begin with the logit equation for $P(A|x, \tau_H = \tau_{CA})$:

[8] As before, we require the same assumption that these individuals don't experience a change in their payoff functions when the policy changes.

[9] Intuitively, this adjustment factor substantially reduces our probabilities P(A) and P(B), because the probabilities of those actions are calculated relative to a much larger denominator. That said, the same scaling factor that is used to adjust our probabilities is also used to adjust the value of $N_{Q,x}$ in Equation B.11. The number of individuals with state variables x_t expands by the same scaling factor and so the adjustments essentially cancel out.

$$P(A|x,\tau_H=\tau_{CA}) = \frac{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))}{\substack{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))+exp(\bar{V}(B,x_i,\tau_H=\tau_{CA}))+exp(\bar{V}(C,x_i,\tau_H=\tau_{CA})) \\ +exp(\bar{V}(D,x_i,\tau_H=\tau_{CA}))+exp(\bar{V}(E,x_i,\tau_H=\tau_{CA}))}}$$

$$= \frac{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))}{\substack{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))+exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))exp(-\theta(f(\tau_{CA}Q_{it})-f(0))exp(-z_i'(\alpha_H-\alpha_Z)) \\ +exp(\bar{V}(C,x_i,\tau_H=\tau_{CA}))+exp(\bar{V}(D,x_i,\tau_H=\tau_{CA}))+exp(\bar{V}(E,x_i,\tau_H=\tau_{CA}))}}$$

Next, we can utilize the fact that we can observe the ratios between the following probabilities $P(A|x,\tau_H=\tau_{CA}), P(B|x,\tau_H=\tau_{CA}), P(C|x,\tau_H=\tau_{CA})$ and $P(D|x,\tau_H=\tau_{CA})$.[10] Based on those figures we can solve for some values of λ_1 and λ_2 such that:

$$\begin{aligned}
P(A|x,\tau_1) = \lambda_1 P(B,x,\tau_H=\tau_{CA}) &\Rightarrow exp(\bar{V}(A,x_i,\tau_H=\tau_{CA})) = \lambda_1 exp(\bar{V}(B,x_i,\tau_H=\tau_{CA})) \\
P(A|x,\tau_1) = \lambda_2 P(C,x,\tau_H=\tau_{CA}) &\Rightarrow exp(\bar{V}(A,x_i,\tau_H=\tau_{CA})) = \lambda_2 exp(\bar{V}(C,x_i,\tau_H=\tau_{CA})) \\
P(A|x,\tau_1) = \lambda_3 P(D,x,\tau_H=\tau_{CA}) &\Rightarrow exp(\bar{V}(A,x_i,\tau_H=\tau_{CA})) = \lambda_3 exp(\bar{V}(D,x_i,\tau_H=\tau_{CA}))
\end{aligned}$$

Inserting those results into the equation for $P(A|x,\tau_H=\tau_{CA})$, we get:

$$P(A|x,\tau_1) = \frac{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))}{\substack{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA})) \\ +exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))(\frac{1}{\lambda_1}) \\ +exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))(\frac{1}{\lambda_2}) \\ +exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}))(\frac{1}{\lambda_3})+exp(\bar{V}(E,x_i,\tau_H=\tau_{CA}))}} \tag{B.12}$$

We can solve that to find that $exp(\bar{V}(A,x_i,\tau_H=\tau_{CA})) = \lambda_4 exp(\bar{V}(E,x_i,\tau_H=\tau_{CA}))$, for some value of λ_4. We then plug that into the equation for $P(A|x,\tau_H=\tau_{CA}-0.01)$ and get:

$$\tag{B.13}$$

$$P(A|x,\tau_H=\tau_{CA}-0.01) = \frac{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}-0.01))}{\substack{exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}-0.01))+ \\ +exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}-0.01))exp(-\theta(f((\tau_{CA}-.01)Q_{it})-f(\tau_{CA}Q_{it}))(\frac{1}{\lambda_1}) \\ +exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}-0.01))exp(-\theta(f((\tau_{CA}-.01)Q_{it})-f(\tau_{CA}Q_{it}))(\frac{1}{\lambda_2}) \\ exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}-0.01))exp(-\theta(f((\tau_{CA}-.01)Q_{it})-f(\tau_{CA}Q_{it}))(\frac{1}{\lambda_3})+exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}-0.01))(\frac{1}{\lambda_4})}}$$

In order to evaluate this expression, we plug in the results from our primary regression. In plugging in our value for θ, we use the coefficient derived from our evaluation of the full primary sample, rather than just individuals residing in California. That said, we estimate our choice probabilities based on those individuals who originate in California. Just as above, evaluating this expression requires several assumptions about the nature of the continuation value functions. First, evaluating this expression requires that across two courses of action the ratio of continuation value functions, $E[V(x_{i,t+2})]$, is not impacted by the change in the high state tax rate. It is this assumption that allows us to re-write: $\bar{V}(B,x_i,\tau_H=\tau_{CA}-.01) = exp(\bar{V}(A,x_i,\tau_H=\tau_{CA}-0.01))exp(-\theta(f((\tau_{CA}-.01)Q_{it})-f(\tau_{CA}Q_{it}))(\frac{1}{\lambda_1})$.[11] Given the set-up here, we should expect that these continuation value functions do not change values after a change in the tax rate. All individuals examined here have already realized their primary capital gains, and so the tax rate on capital gains should have little impact on their decision-making. Second, this calculation requires the assumption that that $exp(\bar{V}(E,x_i,\tau_H=\tau_{CA})) = exp(\bar{V}(E,x_i,\tau_H=(\tau_{CA}-0.01))$. This is a relatively mild assumption as this group is composed of individuals who move a non-zero tax state within two periods and then realize

[10]In the simplest case, we observe this probability directly. In the case where we add more control variables, we need to predict these probabilities to avoid the small cell concerns documented above. In that case, we return to the logit and predict probabilities within tax savings bins, smoothing across incomes, ages and origin destinations.

[11]Our original formula above defined $exp(\bar{V}(A,x_i,\tau_H=\tau_{CA})) = \lambda_1 exp(\bar{V}(B,x_i,\tau_H=\tau_{CA}))$. The following term is included to augment the choice specific value function for Action B, $exp(-\theta(f((\tau_{CA}-.01)Q_{it})-f(\tau_{CA}Q_{it}))$. All other terms in the choice-specific value function are fixed other than the continuation value function term, $E[V(x_{i,t+2})]$.So the ratio holds if the ratio of the continuation value functions is constant.

in that location. These individuals are not faced with any California tax rates and so their payoffs should not be altered by the change in tax rates.

Given the estimate for $P(A|x, \tau_H = \tau_{CA} - 0.01)$, we calculate the following expression to estimate the impact of the policy change on total realization:

$$\sum_Q \sum_x \left(\underbrace{N_{Q,x} Q}\, \underbrace{(P(A|x, \tau_H = \tau_{CA} - 0.01) - P(A|x, \tau_H = \tau_{CA}))} \right)$$

We use values of $N_{Q,x}$and Q based on the set of individuals who originate in California. As noted in the paper, we take the following steps to translate this realization quantity into a fiscal externality: First, we estimate impact of the 1% top tax rate reduction on total realizations by former residents of California. We find that residents of California realize \$13.7 million less in zero-tax states on a yearly basis. After the policy change, the top tax rate in California would be approximately 10% and so a reduction in the top tax rate would increase California revenue by approximately \$1.4 million due to reduced out-migration. Next, we compare that fiscal externality to the mechanical costs of the policy.

Based on the data in our sample, California collected an estimated \$4 billion in yearly capital gains revenue from large realizations between 2005 and 2011.[12] That means that a 1% reduction in top marginal tax rates would have a mechanical cost of \$364 million. Consequently, a \$1.4 million increase in revenue from reduced tax avoidance by out-migrants would offset less than 0.5% of mechanical costs.

The paper notes that this 0.5% may be an upper bound because individuals may not avoid 100% of taxes in their origin state. We explore that possibility using data on the state and local tax (SALT) deduction. In Appendix Figure A.3 below we estimate state income tax liability as a fraction of AGI.

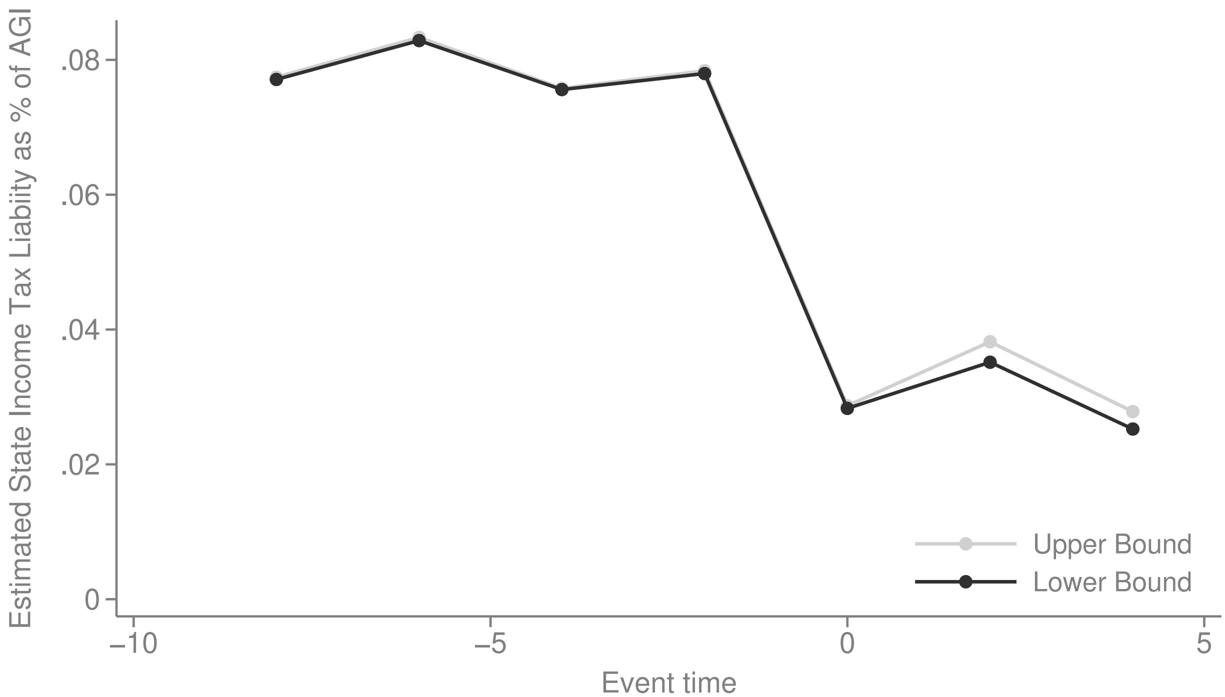
The graph presents results on the SALT deduction for a restricted sample for whom capital gains constituted most of their income. In particular, it is restricted to individuals with non-capital gains income (AGI - capital gains) in the year after realization that is less than 20% of the size of the individual's largest realization. This sample restriction is an attempt to isolate SALT deductions on capital gains income, rather than other sources of income which may be difficult to move across state lines. The graph shows an upper and lower bound because the goal is to use SALT data to isolate the quantity of state income taxes paid after migration. The SALT deduction allows individuals to claim deductions on their property taxes paid as well as their state and local income taxes paid or their sales taxes paid. Filers must choose between claiming the deduction on income taxes or sales taxes.[13] For that reason, it may be insufficient to plot SALT deductions over time as a fraction of AGI. Some taxpayers will claim a deduction on income taxes before their move and switch to claiming the deduction on sales taxes after their move. Those individuals could have avoided state income taxes and it is important not to misinterpret the SALT graphs as proof they failed to avoid state income tax liability. Instead, we plot that the SALT deduction as a fraction of AGI and label that data as an upper bound on the state income tax liability as a share of AGI. For the individuals who chose to deduct their state income taxes, we observe their liability directly. (This makes up approximately 71.8% of individuals after migration.) For the individuals who choose to deduct their local income taxes, we can infer that their state and local income tax liability was less than their sales tax liability. So we know that their state income tax liability is bounded above by their sales tax deduction. Using the same logic, we can estimate a lower bound on state income tax liability by assuming that all individuals who claimed the sales tax deduction has no state income tax liability. Assumption is used in the lower-bound specification in Figure A.3. As seen in the figure, these estimate produce very tight bounds. A relatively small fraction of individuals claim the sales tax deduction (an average of 28.2%) and the mean sales tax deduction is very small as a fraction of total income. Using the drop of deductions as migration, we estimate that migrating to zero tax states results in the avoidance of 60-62% of potential state income taxes on capital gains. This

[12]This data on realizations in our sample is broadly consistently with published public data from the state of California. Data from California suggests that average yearly revenue from capital gains was \$7.2 billion between 2005 and 2011(California Department of Finance, 2015). If 55% of that revenue was collected by the large realizers in our sample, then this would line up with our estimated \$3.56 billion figure. Moreover, data from 2017 in California suggested that the state collected 76% of it's capital gains tax revenue from individuals earning over \$1 million (California Department of Finance, 2019). Given that California's tax code was more progress in 2017 as compared to 2005-2011 and given that this \$7.2 billion figure includes small realizations by individuals with large realizations, these numbers appear to align quite well.

[13]The option to deduct sales taxes was introduced in 2004. For filers who do not report individual purchases, the allowable sales tax deduction is a function of their income. The deduction caps out for incomes above \$300,000, limiting the size of this deduction for high income taxpayers.

confirms to migration to a zero-tax state results in a meaningful decrease in tax liability, but suggests that former residents of high tax states may be unable to move all of their capital gains across state lines.

Figure A.3: Estimated State Income Tax Liability Amongst Individuals who Migrate and then Realize



Notes: This figure uses information from State and Local Tax (SALT) Deductions to plot estimates for state income tax liability as a fraction of AGI. This fraction is plotted in year relative to the time of realization, $t = 0$. The sample is restricted to individuals who migrate to a zero-tax state in the year prior to their realization. The sample is also restricted to individuals with small quantities of non-capital gains income. In particular, it is restricted to individuals with non-capital gains income (AGI - capital gains) in the year after realization that is less than 20% of the size of the individual's largest realization. The upper bound plotted here reports the quantity of deductions claimed on state and local taxes (but not property taxes). The SALT deduction allows individuals to claim deductions on either their state and local incomes taxes or their local sales taxes paid. The upper bound here include state and local deductions claimed on either income or sales tax. The lower bound plotted here reports the quantity of deductions claimed on state and local income taxes alone. As explained below, these two series produce approximate bounds on state income tax liability as a fraction of AGI amongst individuals who migrate to a zero-tax state and then realize.

As noted above, these counterfactual calculations focus on the set of individuals who are observed in a high tax state and realize within 2 periods. That assumption is made for tractability, but we can explore how relaxing that assumption modifies our results. It might be, for example, that some individuals who take Action A would have stayed in the high tax state for two periods and then realized in that state. We can add that course action of action F, into our calculations.[14] It has a choice specific value function,

[14]In this case, we explore the robustness of our results by adding in an individual course of action rather than taking the approach of Section B.V and adding an adjustment term for the total number of individuals with state variable, x_t. In the previous example those alternate courses of action did not have their payoff functions altered by the policy change at hand. That made it possible to make an assumption about the consistency of the payoff function in the counterfactual. In this case, an action such as waiting for 3 periods and then realizing is impacted by the policy change. That makes it essential to incorporate individual course of action and adjust for the change in the payoff function in our counter-factual calculation. It is possible to re-do our analysis with an extension number of additional actions that specify choices for 3 or more periods, but taking such

$\bar{V}(F, x_i, \tau_H = \tau_{CA}) = z_i'\alpha_H + z_i'\alpha_H + z_i'\alpha_H - \theta f(\tau_{CA} Q_{it}) + E[V(x_{i,t+3})]$. Just as before, we can take the relative probability of Action F and Action A to solve for our pre-reform and post-reform probabilities of Action A. There is a change in the payoff function for Action F due to the policy reform. The key assumption in this case is the continuation values for these two actions remain constant after the policy changes. In this case it is the $z'\alpha_H + E[V(x_{i,t+3})]$ for Action F as compared to $E[V(x_{i,t+2})]$ for Action A. As the individuals will have fully realized their gains by period 3 in both cases, this appears to be a reasonable assumption. When making this adjustment we find that the lower of California's tax rate results in \$15.7 million fewer realizations in zero-tax states. This is figure is relative similar to the \$13.7 million estimate found in our baseline specification.

B.VII Reducing Top Tax Rates – Comparisons Across States

After reporting the fiscal externality from reducing California's top tax rate by 1%, we report that the same fiscal externality across all US states. In particular, we note that the maximum fiscal externality is less than 1% of the mechanical costs of reducing top rates. We arrive at that figure by repeating the calculation from Appendix Section B.VI above. In analyzing the state of California we focused on a time period (2005-2011) where there was a unique top tax rate that applied to individuals earning more than \$1 million. (This was California's Mental Health Services Tax.) This surtax made it natural to imagine reducing tax rates by 1pp for individuals in our sample.[15] In analyzing the other states in our sample, we once again consider the impact of reducing the top tax rate by 1pp.[16] We restrict our calculation to the individuals in our sample. So the fiscal externality and the mechanical cost are both calculated based on the taxation of realizations with potential tax savings greater than \$20,000.[17] This means that our fiscal externality is likely a conservative upper bound. In most US states, a reduction in top tax rates would reduce rates on smaller realizations that fall beyond the scope of our analysis. To the extent that those smaller realizations produce smaller behavioral responses, we expect that the fiscal externality rate would fall.

As noted in the text, the fiscal externality rate from a 1pp tax reduction in any given state is approximately proportional to the status quo rate of migration to zero-tax states in advance of realization. This pattern can be seen clearly in Appendix Figure A.4 below.

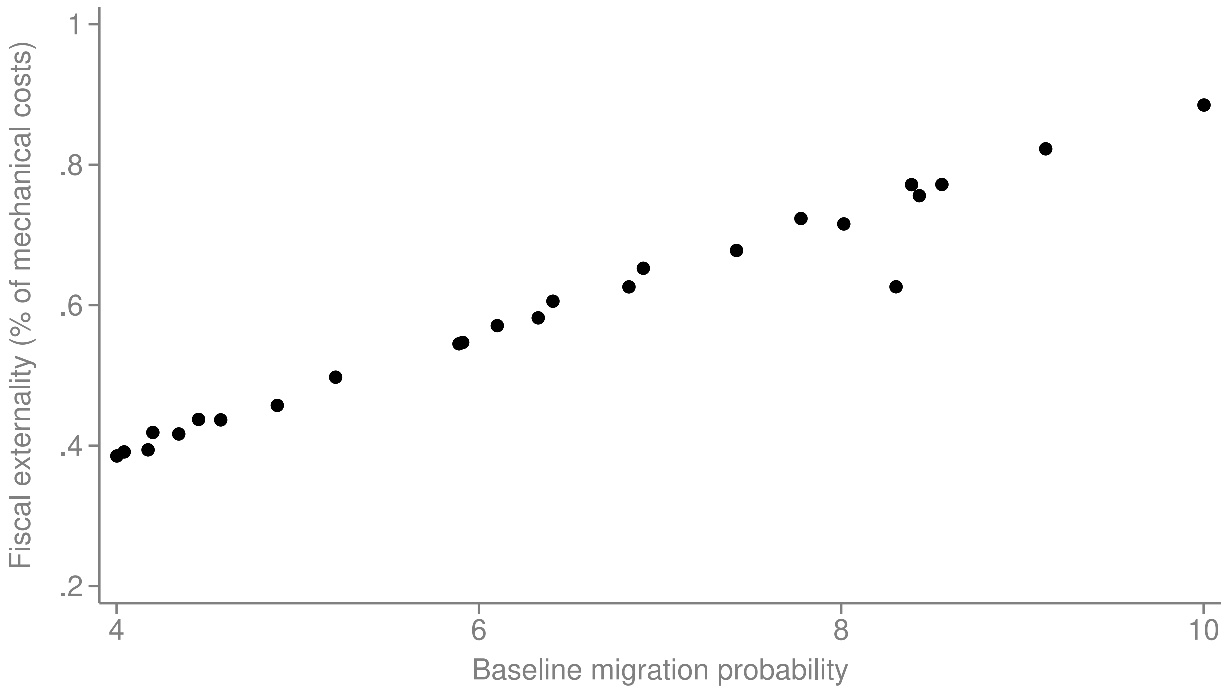
an approach substantially restricts our sample. Instead, we focus on adding the course of action involving realization after 2 years which is likely to be chosen by the highest number of individuals switching away from Action A.

[15] As noted in the paper, 89% of all large capital gains in our sample are dollars above the \$1M AGI threshold.

[16] For consistency, we calculate this on the sample of realizers between 2005-2011. The pattern of our results remain unchanged when calculated over the full sample of available years.

[17] For a state with a 5% tax rate, this corresponds to major realizations in excess of \$600,000.

Figure A.4: Fiscal Externality Across States From Tax Rate Reduction Relative to Baseline Migration to Zero-Tax States



Notes: This figure is based on results from a counterfactual policy change that reduces top tax rates by 1% in each US state. Each dot corresponds to the effect of the policy in one US state over the full sample period. States are plotted along the x-axis based on the status quo probability that individuals with large capital gains migrate to zero-tax state in advance of realization. The probability is reported in realization-weighted terms. States are plotted along the y-axis depending on the fiscal externality produced by the policy as a fraction of the policy's mechanical cost.

In order to understand the reason for this pattern, we can re-examine the probabilities of migration given by Equations B.12 and B.13 above. In particular, we look at the ratio of $\frac{P(A|x,\tau_H=\tau-a)}{P(A|x,\tau_H=\tau)}$ for some small tax change from $\tau_H = \tau$ to $\tau_H = \tau - a$. Simplifying that ratio gives the following:

$$\frac{1+\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\frac{1}{\lambda_3}+(\frac{1}{\lambda_4})}{1+exp(-\theta(f((\tau-a)Q_{it})-f(\tau Q_{it}))(\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\frac{1}{\lambda_3})+(\frac{1}{\lambda_4})}$$

Here, Actions B and C and D represent realizing in one's own state. Those actions occur with very high probabilities. For that reason, the three middle terms in the numerator and denominator are far larger than the remaining two terms.[18] We can therefore re-write this equation as approximately equal to:

$$\begin{aligned}&\approx\frac{\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\frac{1}{\lambda_3}}{exp(-\theta(f((\tau-a)Q_{it})-f(\tau Q_{it}))(\frac{1}{\lambda_1}+\frac{1}{\lambda_2}+\frac{1}{\lambda_3})}\\&\approx 1/exp(-\theta\,(f(\tau Q)-((\tau-a)Q)))\end{aligned}$$

We can then use our functional form for $f(\cdot)$ to again simplify this expression:

[18]More formally $(\frac{1}{\lambda_1})+(\frac{1}{\lambda_2})+(\frac{1}{\lambda_3})$ is much larger than $1+(\frac{1}{\lambda_3})$

$$
\begin{aligned}
\theta\big(f(\tau Q) - f((\tau-a)Q)\big) =&\theta\big(\ln(\tau Q + B) - \ln((\tau-a)Q + B)\big) \\
\approx&\theta\big(\ln(\tau Q) - \ln((\tau-a)Q)\big) \\
=&\ln\left[\Big(\frac{\tau}{\tau-a}\Big)^{\theta}\right]
\end{aligned}
$$

When we exponentiate that expression and take a first-order taylor expansion to get that

$$
\exp\left[\theta\big(f(\tau Q) - f((\tau-a)Q)\big)\right] \approx 1 + \frac{(a)\theta}{\tau}
$$

This tells us that our change in probabilities is approximately proportional to $\frac{(a)\theta}{\tau}$ or that $P(A|x, \tau_H = \tau - a) - P(A|x, \tau_H = \tau) \approx \frac{(a)\theta}{\tau} P(A|x, \tau_H = \tau)$. Our fiscal externality is given by the portion of revenue provided to the origin state by individuals who choose not to realize in a zero-tax location. As a result, it is proportional to $P(A|x, \tau_H = \tau - a) - P(A|x, \tau_H = \tau)$, the change in the probability that, after the tax change, individuals will migrate in advance of realizations. This change in probabilities is approximately equal to $\frac{(a)\theta}{\tau} P(A|x, \tau_H = \tau)$ and so the change in probability is proportional to the baseline rate of migration to zero tax states. This helps explain the clear upward slope found in Appendix Figure A.4 where the fiscal externality is compared to the baseline probability of migrating before realization. The slope of that relationship is governed by θ the coefficient of interest in our regressions.

C Supplemental Robustness

C.I Endogenous quantity decisions

The approach we've taken so far treats realization quantities as fixed. Individuals have some quantity of unrealized capital gains, Q_i, that they choose to realize at one time. This assumption allows us to conduct our analysis using observed realization quantities rather than unobserved quantities of unrealized gains. We consider this a reasonable assumption because the data we observe as consistent with the notion that large realizers are selling major assets such as businesses. (As noted, the major realizations we observe in our sample constitute 82% of dollar-weighted realizations in the 3-year period that contains the major realization event.[19]) Moreover, the data reported in Figure A.1 suggests that observed gains predominately come from business assets rather than other sources such as real estate.) It is worth acknowledging, however, that taxes might change an individuals decision of *how much* to realize, in addition to their decision of *whether* to realize. For example, without any zero-tax migration opportunities, an individual California might sell 90% of their business and hold the remaining 10% until death. Presented with the opportunity to avoid taxes in Florida, that same individual might choose to migrate and realize 100% of their gains. We don't have any way of directly estimating this change in behavior, but we can conduct bounding exercises to show that such behavior is unlikely to substantially impact our results.

The basic intuition of our approach is that we can use alternate comparison groups to account for the possibility that taxes change realization quantities. If, for example, an individual migrates and then realizes \$10 million, the appropriate comparison group may be individuals who realize \$9 million and then migrate. The individual who realizes \$9 million before migration is the type of individual who would have realized \$10 million if they had chosen to switch the relative order of their realization and migration.

We formalize that logic in the following way. Let us assume that a 100% reduction in one's tax rate leads to a 10% increase in realization quantities. In other words, if $Q(0.10) = x$, then $Q(0) = 1.1x$. We let an individual's state variables be $x_{it} = (z_{it}, s_{it-1}, \tau_{s_{t-1}}, Q_i(\tau_{s_{t-1}}))$. When an individual realizes if state s, they realize a quantity $Q_i(\tau_s)$. Using τ_s and our equation for $Q_i(\tau_s)$ we can back out $Q_i(\tau_{s_{t-1}})$.

Working through the same approach from above, we get the following equation:

[19] There is slight variation in realization patterns across income, so this figure is 89% of person-weighted realizations.

$$ln\left[\frac{Pr(s_t = Z, r_t = 0, s_{t+1} = Z, r_{t+1} = 1|x_{it})}{Pr(s_t = H, r_t = 1, s_{t+1} = Z, r_{t+1} = 0|x_{it})}\right] = z'_{it}(\alpha_{H_j} - \alpha_{Z_j}) + (z'_{it} - z'_{it+1})\eta + \theta(f(\tau_{H_j} Q_{it}(\tau_{H_j})) - f(0)) + \Delta\epsilon_t \tag{C.1}$$

This differs from our previous estimation equation in two key ways:

First, when we calculate the left-hand side of our regression equation above, we compare individuals with different observed realization quantities. Both the numerator and the denominator condition on the state variables x_{it}. But for a given value of x_{it} and consequently a given value of $Q_{it}(\tau_{H_j})$, two individuals can have differently sized realizations if they realize in states with different tax rates. The captures the intuition above that when analyzing people who migrate and realize \$10 million, the proper comparison group might be individuals who realize \$9 million and then migrate.

Second, the tax savings term in our regression is now $\theta(f(\tau_{H_j} Q_{it}(\tau_{H_j})) - f(0))$ rather than $\theta(f(\tau_{H_j} Q_{it}(\tau_{Z_j})) - f(0))$. These two terms differ slightly because this new term uses predicted realization quantities as determined by the tax rate τ_{H_j}. For the example given above, this means that we use the \$9 million realization quantity, rather than the \$10 million figure.

In order to implement this regression, we also need to make a slight modification to our sample. Our original sample includes all individuals with potential tax savings over \$20,000. We expand our sample to include individuals with potential tax savings over \$16,667. This ensures that all individuals realizing in zero-tax states have comparison individuals in our sample who realize in high-tax states.

With these modification in mind, we can estimate the coefficient on θ in equation C.1 above. In that case, we can once again examine how the presence of a zero tax opportunity impacts realizations. We find that realizations increase by an estimated \$2.7 billion, compared to the \$2.8 billion in our baseline specification. We also examine how realizations change if zero-tax opportunities increase realizations by 20%. In that case, total realizations change from \$2.8 billion to \$2.4 billion. When we examine the impact on our California realizations specification, we find that the magnitude of the adjustment is slightly larger but qualitatively similar. These results can be found in Appendix Table A.4 below:

Table A.4: Robustness to Endogenous Realization Quantities

	Baseline (1)	10% increase (2)	20% increase (3)
θ Coefficient	0.0965	0.0885	0.0764
	[0.0750, 0.1194]	[0.0659, 0.1117]	[0.0537, 0.1010]
Zero-Tax Realizations	\$2,846m	\$2,660m	\$2,391m
	[2339, 3332]	[2100, 3215]	[1771, 3001]
CA Fiscal Externality	\$13.7m	\$11.3m	\$8.9m
	[10.5, 17.3]	[8.3, 14.9]	[6.1, 12.3]
Income/age controls	X	X	X
Origin controls	X	X	X
Destination controls	X	X	X
Smoothed dependent variable	X	X	X

Notes: Column (1) reproduces the baseline results found in Table 1 of the main text. Column (2) presents the results in the scenario where individuals faced with no state capital gains taxes, increase their realizations by 10%. (The results in Column (1) come from from Equation B.8, whereas the results from Column (2) come from Equation C.1.) Column (3) presents the results in the scenario where individuals faced with no state capital gains taxes, increase their realizations by 20%. In each case, we report our coefficient of interest, θ. This captures the impact of potential tax savings on individual payoffs. Next, we consider a counterfactual where residents of high tax states cannot avoid state capital gains taxes via migration. We compare this counterfactual to the status quo and estimate the effect of the status quo on new realizations in zero-tax states. We report the quantity of new yearly realizations by former residents of high-tax states. Finally, we consider a counterfactual where the state of California reduces its top marginal tax rate by 1%. We report the effect of reduced out-migration on capital gains realizations in zero-tax states. All 95% confidence intervals are obtained using the Bayesian bootstrap.

C.II Incorporation of Discount Rates

In this section, we explore how the incorporation of discounting impacts our results. We show that using a value of β other than 1 has a minimal impact on our results. In order to explore the role of the discount rate, we begin from a modified version of B.7 in B.II above:

$$ln[\frac{Pr(s_t = Z_j, r_t = 0|x_{it})}{Pr(s_t = H_j, r_t = 1|x_{it})}] + \beta[ln\frac{Pr(s_{t+1} = Z_j, r_{t+1} = 1|x_{it+1}(s_t = Z_j, r_t = 0))}{Pr(s_{t+1} = Z_j, r_{t+1} = 0|x_{it+1}(s_t = H_j, r_t = 1))}]$$
$$= z'_{it}(\alpha_{H_j,g} - \alpha_{Z_j,g} + (1-\beta)\eta) + \beta(z'_{it} - z'_{it+1})\eta + \theta\left(f(\tau_H Q_{it}) - f(\beta\tau_Z Q_{it} * (1+R))\right) + \Delta\epsilon_t$$

In this case we re-arranged the first term on the right hand side so that it captures both the preference for the high-tax state over the zero-tax state and for the impact of delaying moving costs for a year. (Intuitively, the choice to realize before migration could be driven by a preference to reside in the high tax location for an additional year or by a preference to delay moving by a year.) Once again the second term captures the impact of time-varying demographic variables on moving costs and the third term captures the payoff associated with additional tax savings. We have also moved the discount term, β, inside the function $f(\cdot)$ capturing the payoff to tax savings. The idea here is that individuals value the discounted flow of future tax savings, rather than discount the payoff from the flow of future tax savings.

In order to specify our final regression equation we rearrange the left-hand side of this equation. We write the left-hand side as $\beta ln\left[\frac{Pr(s_t=Z_j,r_t=0,s_{t+1}=Z_j,r_{t+1}=1|x_{it})}{Pr(s_t=H_j,r_t=1,s_{t+1}=Z_j,r_{t+1}=0|x_{it})}\right] + ln[\frac{Pr(s_t=Z_j,r_t=0|x_{it})}{Pr(s_t=H_j,r_t=1|x_{it})}]$. Putting those together, we get the following:

$$
\beta ln\left[\frac{Pr(s_t = Z_j, r_t = 0, s_{t+1} = Z_j, r_{t+1} = 1|x_{it})}{Pr(s_t = H_j, r_t = 1, s_{t+1} = Z_j, r_{t+1} = 0|x_{it})}\right] + (1-\beta)ln[\frac{Pr(s_t = Z_j, r_t = 0|x_{it})}{Pr(s_t = H_j, r_t = 1|x_{it})}]
$$
$$
z'_{it}(\alpha_{H_j,g} - \alpha_{Z_j,g} + (1-\beta)\eta) + \beta(z'_{it} - z'_{it+1})\eta + \theta\left(f(\tau_H Q_{it}) - f(\beta\tau_Z Q_{it} * (1+R))\right) + \Delta\epsilon_t \tag{C.2}
$$

We estimate this equation into order to solve for our coefficient of interest, θ. The first term on the left-hand side is calculated in the same manner as our primary regression in C.I above. It is simply multiplied by our various calibrated values of β. Inspired by the use of a 3% real rate of interest we begin with a value of $\beta = 0.94$. This is because we examine individuals who realize within a four-year window after realization. The 0.94 value accounts for the fact that the median individual delays by 2 years. We also examine alternative values of 0.9 and 0.97. The second term on the right-hand side is estimated by comparing the set of individuals who realize in their home state to the set of individuals who migrate to a zero-tax state without having realized their gains.[20]

Once we solve for the value of θ, we can examine how the presence of a zero tax opportunity impacts realizations. For a discount factor of $\beta = 0.94$, we find that realizations increase by an estimated $2.7 billion, compared to the $2.8 billion in our baseline specification. We also examine how realizations change for discount factors of 0.90 and 0.97. In that case, total realizations change to $2.6 billion and $2.7 billion, respectively. When we examine the impact on our California realizations specification, we find that the magnitude of the effect changes from $13.7 million to $12.1 million after the use of a discount factor of $\beta = 0.94$. These results can be found in Appendix Table below:

[20] The key to calculating the size of this second group is to determine the set of individuals with the right set of state variables, x_{it}. We want individuals with a given quantity of unrealized gains. For that reason, this group includes any individuals who realize that quantity in the years after their migration. This is a different group of individuals than those in the numerator of our first left-hand side term because this includes individuals who realize multiple periods after their move. Here, we expand our sample using the same approach adopted in Section B.V where we look at the total population of individuals realizing from 1999 onward. In particular, the numerator is composed of individuals who resided in a high tax state in 1998, migrated to a zero tax state in 1999-2002 and realized at some point between 2003 and 2019. The denominator is composed of individuals who resided in a high tax state in 1998 and realized in that state from 1999-2002. Ideally, we would also want to incorporate individuals who have such unrealized gains but hold those gains until death. Unfortunately, those gains are unobserved and so those individuals cannot be incorporated here. That said, this ratio is scaled by $(1-\beta)$ and so this second term on the left-hand side of the equation is small relative to the first term. Consequently, changes to the denominator have a small impact on the total value of left hand side of the equation.

Table A.5: Robustness to Various Discount Factors

	Baseline: $\beta = 1$ (1)	$\beta = 0.97$ (2)	$\beta = 0.94$ (3)	$\beta = 0.9$ (4)
θ Coefficient	0.0965	0.0915	0.0909	0.0871
	[0.0750, 0.1194]	[0.0700, 0.1155]	[0.0693, 0.1142]	[0.0643, 0.1104]
Zero-Tax Realizations	$2,846m	$2,729m	$2,714m	$2,633m
	[2339, 3332]	[2207, 3235]	[2178, 3229]	[2072, 3196]
CA Fiscal Externality	$13.7m	$12.6m	$12.1m	$11.1m
	[10.5, 17.3]	[9.3, 16.3]	[9.0, 15.7]	[8.1, 14.5]
Income/age controls	X	X	X	X
Origin controls	X	X	X	X
Destination controls	X	X	X	X
Smoothed dependent variable	X	X	X	X

Notes: Column (1) reproduces the baseline results found in Table 1 in the main text. Column (2) presents the results assuming a per-period discount factor of 0.9. The coefficient of interest θ is estimated in Equation C.2. Section C.II explains the modifications to the primary regression equation to incorporating this discounting. This estimate of θ captures the impact of potential tax savings on individual payoffs. Next, we consider a counterfactual where residents of high-tax states cannot avoid state capital gains taxes via migration. We compare this counterfactual to the status quo and estimate the effect of the status quo on new realizations in zero-tax states. We report the quantity of new yearly realizations by former residents of high-tax states. Finally, we consider a counterfactual where the state of California reduces its top marginal tax rate by 1%. We report the effect of reduced out-migration on capital gains realizations in zero-tax states. Column (3) repeats this analysis using a per-period discount factor of 0.94 while Column (4) uses a per-period discount factor of 0.97. All 95% confidence intervals are obtained using the Bayesian bootstrap.

C.III Estimates with a Treatment-Control Framework

In this section we dispense with the dynamic discrete choice model explained in Appendix B and examine how our results would change if we adopted a simple linear probability model in a treatment-control framework. In this case we measure the probability of migrating in advance of realization in the following manner:

$$P_{ioyt} = \alpha_o + \eta_y + \beta Q_i + \theta(T * Q_i) + \varepsilon \tag{C.3}$$

- P_{ioyt} is an indicator for whether individual i migrates to a zero tax state in period t.
- α_o is an origin state fixed effect
- η_y is an income-bin fixed effect
- β captures the impact of realization quantities on migration
- T is the tax differential between high tax states and zero-tax states
- θ is our coefficient of interest

In this case, our goal is to estimate θ. This approach requires making an assumption about the presence of a separate treatment and control group. In this case, we assume that individuals who migrate before realization are the treatment group and those who realize before migration are the control group. (This is not precisely correct membership in the control group is endogenous. It is only composed of those who did not choose to migrate before realizing. With that assumption in place, we can estimate how the probability of migration changes with tax savings.

That allows us to estimate our first counter-factual with the following expression:

$$\sum\left(N_q Q\hat{\theta}(\bar{T_{hi}} - 0)Q\right)$$

Similar to Equation 5 in the paper, this expression evaluates how changes in the tax rate change the probability of migrating and then multiplies that figure by the quantity of individuals realizing in each realization bin and the average realization size in that bin. Using this specification, we estimate that the status quo results in $2.7 billion in additional yearly realizations relative to the counter-factual. This estimate is very similar to the $2.8 billion estimate produced by our full dynamic discrete choice model. When we examine the impact on California realizations the effect size rises from $13.7 million to $29.5 million but the fiscal externality remains less than 1% of the mechanical cost of the policy.

C.IV Multi-Year Realization Events

In our primary analysis we classify individuals by the size of their largest realization event. As we note in Section II, for individuals in our sample, this largest gain represents 89% of their gains in a 3-year window and 83% of their gains in a 5-year window. There is slight variation in this number across realization distribution but this 3-year figure is still 82% in dollar-weighted terms. This suggests that the our results are primarily driven by the one-year realizations we observe. After all, if we assumed that our treatment effects remained constant but we were underestimating the quantity of realizations moved across state lines, our counterfactual estimates would change minimally. If for example, our Q in the following equation was 82% of it's true value, $\sum_Q\sum_x\left(\underbrace{N_{Q,x}Q}\underbrace{(P(A|x,\tau_Z=\tau_H)-P(A|x,\tau_Z=0))}\right)$, then zero tax opportunities would result in $3.4 billion in yearly realizations rather than $2.8 billion.

In order to investigate this issue more formally, we recalculate our dynamic discrete choice model assuming that individuals are motivated by 3-year realization events rather than 1-year realization events. In other words, we our state variable $\tau_{s_{t-1}}Q_i$ classifies individuals by their 3-year tax savings rather than their one year savings. (For simplicity we still determine the relative timing of realizations and migrations based on the year of the greatest realization.) Using this approach we get a value of θ equal to 0.0924 and estimate that zero-tax opportunities lead to $3.4 billion in new realizations, rather than the $2.8 billion in our baseline specification. These results can be found in Appendix Table A.6 below:

Table A.6: Additional Robustness of Primary Results

	Baseline (1)	**3-year** (2)	**Nonzero tax migration** (3)	**Expanded choice set** (4)
θ Coefficient	0.0965	0.0924	-0.0601	0.0965
	[0.0750, 0.1194]	[0.0714, 0.1149]	[-0.0964, -0.0228]	[0.0760, 0.1199]
Zero-Tax Realizations	$2,846m	$3,374m		$2,912m
	[2339, 3332]	[2763, 3987]		[2422, 3430]
CA Fiscal Externality	$13.7m	$15.5m		$15.7m
	[10.5, 17.3]	[11.7, 19.8]		[12.3, 19.7]
Income/age controls	X	X	X	X
Origin controls	X	X	X	X
Destination controls	X	X	X	X
Smoothed dependent variable	X	X	X	X

Notes: Column (1) reproduces the baseline results found in Table 1 in the main text. Column (2) presents the results assuming that individuals are motivated by 3-year realization events rather than 1-year events. In each case, we report our coefficient of interest, θ. This captures the impact of potential tax savings on individual payoffs. Next, we consider a counterfactual where residents of high tax states cannot avoid state capital gains taxes via migration. We compare this counterfactual to the status quo and estimate the effect of the status quo on new realizations in zero-tax states. We report the quantity of new yearly realizations by former residents of high-tax states. Finally, we consider a counterfactual where the state of California reduces its top marginal tax rate by 1%. We report the effect of reduced out-migration on capital gains realizations in zero-tax states. Column (3) considers the impact of migration to non-zero tax states. It presents the results results of a scenario where a separate coefficient of interest, θ_2, is estimated as part of the payoff function of individuals migrating from high-tax to low-tax states. That coefficient estimated using Equation C.5. Column (4) examines the results in the case where the sample is expanded to incorporate individuals who don't realize within the two-period window used in the primary sample. As explained in Section B.V, the probabilities used in our first counterfactual are scaled so that rates of migration and realization are calculated relative to the total number of individuals in high tax states who realize between 2003 and 2018. As explained in Section B.VI, the courses of action in our second counterfactual are altered to include individuals who reside in a high-tax state for 3 periods before realizing. All 95% confidence intervals are obtained using the Bayesian bootstrap.

C.V Migration to Non-Zero Tax States

In our primary analysis, we focus on migration to zero-tax states. The choice to focus on this group of moves is driven by the data. Figure IV in the main text shows rates of migration relative to realization for individuals moving to low, but non-zero tax states. As discussed in the text, the data do not point to clear tax avoidance behavior amongst those migrating to non-zero tax states. Migration in advance of realization appears to decline with the size of tax savings. This suggests that individuals with large potential gains have a desire to substitute away from non-zero tax locations (and potentially toward zero-tax locations). In order to formally confirm these patterns we re-estimate our dynamic discrete choice model with a new set of payoff functions for migrants to non-zero tax states. We consider those individuals as having a payoff function[21]:

$$\pi_{it}(s,r) = z'_{it}(\alpha_{2,s,g} + \eta_2 m_{it}(s_{it}, s_{it-1})) - \theta_2 f(\tau_{s_{t-1}} Q_{it}) r_{it} \tag{C.4}$$

Here, θ_2 is our new coefficient of interest and it can differ from our original value of θ. Solving the dynamic discrete choice model we can estimate the coefficient from the following regression:

$$\begin{aligned} ln\left[\frac{Pr(s_t = M_j, r_t = 0, s_{t+1} = M_j, r_{t+1} = 1 | x_{it})}{Pr(s_t = H_j, r_t = 1, s_{t+1} = M_j, r_{t+1} = 0 | x_{it})}\right] =& z'_{it}(\alpha_{H_j,g} - \alpha_{M_j,g}) + (z'_{it} - z'_{it+1})\eta \\ &+ \theta_2 \left(f(\tau_H Q_{it}) - f(\tau_M Q_{it})\right) + \Delta\epsilon_t \end{aligned} \tag{C.5}$$

[21] We also avoid this approach in our primary specification as it requires assigning different courses of action different payoff functions. While it is true that different groups of individuals likely have different payoff functions, the adjustment made herein is merely a crude way to capture that pattern and demonstrate that migration to non-zero tax states have a minimal impact on our results.

Here, $s_t = M_j$ represents the choice to go to a non-zero tax states. We report the results in Appendix Table A.6. Our previous coefficient for $\theta = .0965$ and here we find a coefficient of -0.06. The presence of a sizable negative coefficient is consistent with the presence of a substitution effect, individuals who might have migrated to lower-tax states choose to take advantage of the tax savings in zero-tax states and migrate there instead.

It is worth noting that these results have important implications for the initial value of θ estimated in equation 2 in the paper. That analysis estimates θ by comparing those who migrate before realization with those who realize before migration. It compares across individuals with different realization quantities because those individuals have different amounts of potential tax savings. That analysis implicitly assumes that when comparing across individuals we are changing potential tax savings while holding all else fixed. In truth, individuals who have higher potential tax savings when migrating to zero-tax states also have higher potential tax savings when migrating to lower-tax states. If we saw large numbers of individuals moving to lower-tax states, that could pose a problem for our estimate of θ. Our log odds ratio would could be impacted by the relative desirability of migration to the zero-tax versus low-tax state. In practice, rates of migration to low-tax states are quite small relative to rates of migration to zero-tax states. Even more importantly, the presence of a negative coefficient in Equation C.5 above suggests that migration to low-tax states is not meaningfully impacting our results. When looking at individuals who migrate to low-tax states before or after realization, we are seeing that the effect of tax savings in zero-tax states dominates the effect of tax savings in low-tax states. That is how we get a coefficient that is negative but of similar absolute value of our initial estimate of θ. In other words, tax motivated migration and realization decisions are driven by the savings in zero-tax states, rather than low-tax states.

References

Almagro, M. and T. Domínguez-Iino (2021). Location sorting and endogenous amenities: Evidence from amsterdam.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

California Department of Finance (2015). Governor's budget summary, 2014-2015. `http://www.ebudget.ca.gov/2014-15/pdf/BudgetSummary/FullBudgetSummary.pdf`.

California Department of Finance (2019). California department of finance fall 2019 report. `https://data.ftb.ca.gov/stories/s/Other-Tax-Data/sgxz-494k#department-of-finance-exhibits`.

Diamond, R., T. McQuade, and F. Qian (2017). The effects of rent control expansion on tenants, landlords, and inequality: Evidence from san francisco.

Feenberg, D. R. and E. Coutts (2018). Maximum state income tax rates. `https://users.nber.org/~taxsim/state-rates/`.

Kalouptsidi, M., P. T. Scott, and E. Souza-Rodrigues (2020). Linear IV regression estimators for structural dynamic discrete choice models. *Journal of Econometrics*.

Scott, P. (2013). Dynamic discrete choice estimation of agricultural land use.