

# The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists

By ENRICO MORETTI AND DANIEL J. WILSON

## Online Appendix

In Section 2.2, we described the general construction of our tax variables. Here we provide further detail.

We obtain business tax and credit rates by state and year for corporate income taxes, investment tax credits, and research and development tax credits from Chirinko & Wilson (2008), Wilson (2009), and Moretti & Wilson (2014). We use the “effective” state corporate income tax rate, which adjusts the statutory tax rate for deductibility of state taxes on federal corporate tax returns and, in some states, vice versa. It is a function of the state statutory tax rate, the federal statutory tax rate, and the share of federal taxes that are deductible from state corporate income for each state. (See Chirinko & Wilson 2008 for further details.)

To construct the average and marginal individual income tax rates for a hypothetical taxpayer, at a specified income level, in each state and year, we used the National Bureau of Economic Research (NBER) Internet TaxSim calculator. We first obtained data, for each year from 1977 to 2010, on the 95<sup>th</sup>, 99<sup>th</sup>, and 99.9<sup>th</sup> percentiles of the national income distribution (separately for salary and capital gains) from the World Top Incomes Database (Alvaredo, Atkinson, Piketty & Saez 2013) and data on the 50<sup>th</sup> percentile from the Congressional Budget Office (CBO). These nominal income levels then became inputs – as primary wage earnings and long-term capital gains – into the NBER’s Internet TaxSim calculator. Other inputs were the following assumptions: The taxpayer was a married joint filer, less than 65 years old, had zero dependent exemptions, zero childcare expenses, no other sources of income, and zero itemized deductions other than the deduction for state income tax payments (which is calculated by TaxSim). In other words, aside from primary wage earnings and long-term capital gains, all other inputs into TaxSim were 0. We used the stata `taxsim9.ado` interface (available at <http://users.nber.org/~taxsim/taxsim-calc9/index.html>) to run the Internet TaxSim program (a Fortran program).

The Internet TaxSim is a program that calculates federal and state tax variables for a hypothetical taxpayer from a given state in a given year. The program is coded based on information from federal and state tax forms and “summaries published by the Commerce Clearing House, the Advisory Commission on Intergovernmental Relations (ACIR), and the Tax

Foundation” (Feenberg and Coutts, 1993). As described at [www.nber.org/~taxsim](http://www.nber.org/~taxsim), “[a]lthough state tax regimes differ significantly from one another, most share the basic structure of the federal tax. That is, deductions and exemptions are subtracted from adjusted gross income to obtain taxable income. A schedule converts taxable income to income before credits, from which a variety of credits, some refundable, are subtracted.” These specifics of state tax codes, along with a similar tax schedule from the federal tax code, are coded into the TaxSim program.

TaxSim returns the following variables (*inter alia*) for each state and year: federal income tax liability, federal adjusted gross income (AGI), federal marginal tax rate, state income tax liability, state adjusted gross income, and the state marginal tax rate. Note that these federal tax variables can differ by state because of the fact that state tax payments are deductible from ordinary federal AGI for determining federal taxable income, though state taxes are not deductible for the alternative minimum tax (AMT). We calculate the average tax rate (ATR) for an individual in each state\*year as the sum of state and federal tax liability divided by gross income.

As mentioned in the paper, we also constructed an alternative measure of the individual income ATR based on (1) data from a large sample of actual federal tax returns from the IRS Statistics on Income (SOI) that are housed at the NBER and (2) the NBER’s non-public internal TaxSim program, which is more detailed than the public Internet TaxSim.<sup>1</sup> These yearly SOI samples are not public but are available for use by NBER research affiliates. For this measure, we started with the same income levels for primary wage earnings and long-term capital gains as described above. For each year from 1977 to 2010 and each income level (50<sup>th</sup>, 95<sup>th</sup>, 99<sup>th</sup>, and 99.9<sup>th</sup> percentiles of national income distribution), we then had a subsample of the full SOI sample of federal tax returns extracted for taxpayers with total wage and long-term capital gains income within +/- 1% of that income level. This resulted in subsamples with roughly 150 to 750 tax returns. For each subsample, the average value was calculated for each variable that is used as an input into the TaxSim programs. These inputs include the number of dependent exemptions, rental and other income, charitable donations, mortgage interest payments, and various other credits and deductions.

Because these tax returns do not identify the state of residence of the taxpayer, one cannot generate these averages/inputs by state. However, there are two state-specific adjustments that we make in order to generate 50 state-specific versions of each yearly income-threshold vector of inputs. First, following Bakija and Gentry (2014), we apply a state-specific multiplier to the

---

<sup>1</sup> We thank Dan Feenberg at the NBER for extracting subsamples, for relevant income levels, from the SOI taxpayer samples and using these to calculate average values of each tax input variable used by TaxSim.

national average property tax deduction. The numerator of this multiplier is calculated as the total statewide property tax revenues divided by state personal income. Property tax revenues data come from the Census Bureau's Survey of State and Local Government Finances; personal income data come from the BEA. The denominator is the population-weighted average of this ratio across the 50 states. Thus, the multiplier measures the effective property tax rate for the average individual in the state relative to the effective property tax rate for the average individual in the nation. The multiplier for state  $s$  is multiplied by the national average property tax deduction (for a given income threshold and year) to obtain an estimate of the average property tax deduction for the hypothetical taxpayer in state  $s$ .

The second state-specific adjustment is done internally by the TaxSim program. For a given state, the program calculates the state income tax liability based on the inputs described above (i.e., the national averages for each input variable and the state-specific property tax deduction estimate). This state income tax liability is then used as the input into the TaxSim program for the state income tax deduction. In some states, federal tax payments are deductible on the state tax return. In these cases, the program iterates between the state and federal calculations until they converge. The TaxSim program outputs estimates of AGI and state and federal tax liabilities, from which we calculate the ATR as the sum of state and federal tax liabilities divided by gross income.

This alternative approach of using averages of tax input variables based on actual samples of tax returns, rather than simply assuming zeros for various itemized deductions and for income other than wage earnings and long-term capital gains, and using the more detailed internal NBER TaxSim program, rather than the Internet TaxSim program, likely yields more accurate measures of the ATR. However, the approach requires access to the NBER's non-public SOI tax return samples and NBER's non-public internal TaxSim program, which generally is only available to NBER affiliates. Thus, for the sake of replicability, we opted to use ATR constructed as described above, involving only publicly available data and the public version of TaxSim, for the results in the paper. We note, however, that virtually all of the results in the paper are quite similar using this alternative measure of the ATR.

## References

- Bakija, Jon M., and William M. Gentry. "Capital Gains Taxes and Realizations: Evidence from a Long Panel of State-Level Data." Unpublished Manuscript, Williams College.
- Chirinko, Robert, and Daniel J. Wilson. "State Investment Tax Incentives: A Zero-Sum Game?" *Journal of Public Economics* 92(12), Dec. 2008, pp. 2362-2384.

Moretti, E, and D. Wilson (2014) "State Incentives for Innovation, Star Scientists and Jobs: Evidence from Biotech", *The Journal of Urban Economics*, forthcoming

Feenberg, Daniel Richard, and Elizabeth Coutts, An Introduction to the TAXSIM Model, *Journal of Policy Analysis and Management* vol 12 no 1, Winter 1993, pages 189-194.

Wilson, Daniel J. (2009). "Beggar Thy Neighbor? The In-State, Out-of-State, and Aggregate Effects of R&D Tax Credits." *Review of Economics and Statistics* 91(2), 431-436.

Table Appendix 1: Summary Statistics over 51 states-  
Star Population, 99th Perc. ATR, and Corp. Income Tax

	Pop Star 95 (1)	Pop Star 95 (2)	99th Perc. ATR (3)	99th Perc. ATR (4)	CIT (5)	CIT (6)
	2006 value	1977-2006 %-change	2006 value	1977-2006 %-change	2006 value	1977-2006 %-change
Alabama	26.0000	-0.5667	0.2899	-0.1421	0.0650	0.3000
Alaska	0.0000	-1.0000	0.2541	-0.3094	0.0940	0.0000
Arizona	267.0000	1.3421	0.2947	-0.1708	0.0697	-0.3362
Arkansas	10.0000	0.0000	0.3171	-0.1036	0.0650	0.0833
California	3869.0000	1.8852	0.3348	-0.1048	0.0884	-0.0178
Colorado	192.0000	0.3617	0.2989	-0.1500	0.0463	-0.0740
Connecticut	290.0000	-0.1667	0.3031	-0.0349	0.0750	-0.2500
Delaware	77.0000	-0.4539	0.3093	-0.1924	0.0750	0.0000
DC	7.0000	-0.3000	0.3359	-0.1090	0.0998	0.0081
Florida	238.0000	0.1174	0.2541	-0.1833	0.0550	0.1000
Georgia	174.0000	1.2308	0.3122	-0.1313	0.0600	0.0000
Hawaii	10.0000	0.4286	0.3270	-0.1280	0.0640	-0.0054
Idaho	219.0000	20.9000	0.3281	-0.1073	0.0760	0.1692
Illinois	419.0000	-0.5117	0.2838	-0.1536	0.0730	0.8250
Indiana	129.0000	-0.4921	0.2879	-0.1264	0.0850	0.4167
Iowa	82.0000	0.1549	0.3160	-0.1196	0.1200	0.2000
Kansas	41.0000	-0.2115	0.3145	-0.0871	0.0735	0.0889
Kentucky	60.0000	-0.2208	0.3120	-0.0981	0.0700	0.2069
Louisiana	24.0000	-0.6800	0.2946	-0.1087	0.0800	0.0000
Maine	14.0000	0.0769	0.3326	-0.0961	0.0893	0.2757
Maryland	154.0000	-0.2261	0.3003	-0.1518	0.0700	0.0000
Massachusetts	645.0000	0.2697	0.3058	-0.1613	0.0950	0.0000
Michigan	470.0000	-0.1826	0.2925	-0.1711	0.0190	-0.1915
Minnesota	593.0000	2.1711	0.3264	-0.1377	0.0980	-0.1833
Mississippi	12.0000	-0.0769	0.3013	-0.1022	0.0500	0.2500
Missouri	99.0000	-0.3694	0.3093	-0.0907	0.0625	0.2500
Montana	5.0000	-0.2857	0.3171	-0.1336	0.0675	0.0000
Nebraska	31.0000	-0.1389	0.3229	-0.1145	0.0781	0.5778
Nevada	44.0000	2.1429	0.2541	-0.1833	0.0000	-1.0000
New Hampshire	66.0000	0.9412	0.2541	-0.1833	0.0925	0.1563
New Jersey	605.0000	-0.4271	0.3073	-0.0792	0.0900	0.2000
New Mexico	40.0000	1.2222	0.3015	-0.1464	0.0760	0.5200
New York	1265.0000	0.2199	0.3200	-0.1843	0.0750	-0.2500
North Carolina	327.0000	1.5349	0.3301	-0.1027	0.0690	0.1500
North Dakota	5.0000	0.2500	0.2943	-0.1826	0.0700	0.1667
Ohio	398.0000	-0.4033	0.3141	-0.0643	0.0850	0.0625
Oklahoma	86.0000	-0.5635	0.3076	-0.1345	0.0600	0.5000
Oregon	291.0000	4.1964	0.3404	-0.1246	0.0660	-0.0571
Pennsylvania	502.0000	-0.3597	0.2848	-0.1399	0.0999	0.0516
Rhode Island	33.0000	0.1000	0.3212	-0.1230	0.0900	0.1250
South Carolina	59.0000	0.0727	0.3170	-0.1350	0.0500	-0.1667
South Dakota	2.0000	0.0000	0.2541	-0.1833	0.0000	-1.0000
Tennessee	91.0000	0.0706	0.2541	-0.1833	0.0650	0.0833
Texas	956.0000	1.0297	0.2541	-0.1833	0.0450	0.0588
Utah	84.0000	0.7500	0.3122	-0.1156	0.0500	0.2500
Vermont	163.0000	7.5789	0.3230	-0.1380	0.0975	0.3000
Virginia	95.0000	-0.2857	0.3092	-0.1381	0.0600	0.0000
Washington	535.0000	4.4592	0.2541	-0.1833	0.0000	-1.0000
West Virginia	4.0000	-0.8750	0.3156	-0.1051	0.0900	0.5000
Wisconsin	216.0000	0.0588	0.3162	-0.1666	0.0790	0.0000
Wyoming	3.0000	-0.4000	0.2541	-0.1833	0.0000	-1.0000

Table Appendix 2: CIT Regressed on Annual Stock Return of Top Patenters

	No Lags No State FEs	No Lags Year, State FEs	1-year Lag No State FEs	1-year Lag Year, State FEs	1-4 years Lag No State FEs	1-4 years Lag Year, State FEs
Annual Stock Return Growth	-0.0002 (0.0009)	0.0002 (0.0002)	-0.0002 (0.0009)	0.0002 (0.0002)	-0.0002 (0.0009)	0.0001 (0.0002)
Annual Stock Return Growth- 1 year lag			-0.0002 (0.0009)	0.0000 (0.0002)	-0.0002 (0.0009)	0.0000 (0.0002)
Annual Stock Return Growth- 2 year lag					-0.0002 (0.0009)	0.0001 (0.0002)
Annual Stock Return Growth- 3 year lag					-0.0002 (0.0009)	0.0000 (0.0002)
Annual Stock Return Growth- 4 year lag					-0.0003 (0.0009)	0.0000 (0.0002)
Joint Sig. p-value			.7988	.9202	.9934	.9964

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Each column is a separate regression. Sample consists of 1527 observations over 51 states and 30 years. All regressions include year fixed effects.

Table Appendix 3: ATR Regressed on Annual Stock Return of Top Patenters

	No Lags No State FEs	No Lags Year, State FEs	1-year Lag No State FEs	1-year Lag Year, State FEs	1-4 years Lag No State FEs	1-4 years Lag Year, State FEs
Annual Stock Return Growth	0.0008 (0.0007)	0.0000 (0.0002)	0.0009 (0.0007)	0.0000 (0.0002)	0.0008 (0.0007)	0.0000 (0.0002)
Annual Stock Return Growth- 1 year lag			0.0007 (0.0007)	-0.0000 (0.0002)	0.0007 (0.0007)	-0.0000 (0.0002)
Annual Stock Return Growth- 2 year lag					0.0008 (0.0007)	0.0000 (0.0002)
Annual Stock Return Growth- 3 year lag					0.0009 (0.0007)	0.0001 (0.0002)
Annual Stock Return Growth- 4 year lag					0.0009 (0.0007)	0.0002 (0.0002)
Joint Sig. p-value			.3078	.9845	.2787	.7465

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Each column is a separate regression. Sample consists of 1527 observations over 51 states and 30 years. All regressions include year fixed effects.

Table Appendix 4: ITC Regressed on Annual Stock Return of Top Patenters

	No Lags No State FEs	No Lags Year, State FEs	1-year Lag No State FEs	1-year Lag Year, State FEs	1-4 years Lag No State FEs	1-4 years Lag Year, State FEs
Annual Stock Return Growth	0.0002 (0.0007)	0.0005 (0.0005)	0.0002 (0.0007)	0.0006 (0.0005)	0.0002 (0.0007)	0.0008 (0.0005)
Annual Stock Return Growth- 1 year lag			0.0000 (0.0007)	0.0005 (0.0005)	-0.0000 (0.0008)	0.0007 (0.0005)
Annual Stock Return Growth- 2 year lag					0.0000 (0.0008)	0.0006 (0.0005)
Annual Stock Return Growth- 3 year lag					-0.0001 (0.0008)	0.0006 (0.0005)
Annual Stock Return Growth- 4 year lag					-0.0004 (0.0008)	0.0002 (0.0005)
Joint Sig. p-value			.9811	.3425	.9898	.3562

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Each column is a separate regression. Sample consists of 1527 observations over 51 states and 30 years. All regressions include year fixed effects.

Table Appendix 5: RD Regressed on Annual Stock Return of Top Patenters

	No Lags No State FEs	No Lags Year, State FEs	1-year Lag No State FEs	1-year Lag Year, State FEs	1-4 years Lag No State FEs	1-4 years Lag Year, State FEs
Annual Stock Return Growth	0.0008 (0.0014)	0.0011 (0.0011)	0.0008 (0.0014)	0.0012 (0.0011)	0.0008 (0.0015)	0.0012 (0.0011)
Annual Stock Return Growth- 1 year lag			0.0004 (0.0014)	0.0007 (0.0011)	0.0004 (0.0015)	0.0007 (0.0011)
Annual Stock Return Growth- 2 year lag					0.0001 (0.0015)	0.0005 (0.0011)
Annual Stock Return Growth- 3 year lag					0.0000 (0.0015)	0.0004 (0.0011)
Annual Stock Return Growth- 4 year lag					0.0000 (0.0015)	0.0003 (0.0011)
Joint Sig. p-value			.7771	.5395	.9993	.9469

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Each column is a separate regression. Sample consists of 1527 observations over 51 states and 30 years. All regressions include year fixed effects.

Table Appendix 6: The Effect of Net-of-Tax Rates on Outmigration of Star Scientists: Baseline Models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MTR, 99th Perc. (1 - mtr)	1.5886* (0.8318)	2.4464*** (0.5660)	2.3671*** (0.5851)	1.7007*** (0.5240)	2.3202*** (0.5774)	2.3621*** (0.5533)	1.6966*** (0.6045)	3.0825*** (0.9437)
State CIT Rate (1 - cit)	-3.5571*** (0.9292)	1.7468** (0.6957)	1.6077** (0.6914)	1.3334** (0.6011)	0.9661 (0.7297)	1.6550** (0.6456)	1.8861*** (0.6410)	1.6137 (1.1929)
State ITC (1 + itc)	5.1639*** (1.4461)	1.7425*** (0.4532)	1.6678*** (0.5352)	1.6131*** (0.5646)	1.7213*** (0.4949)	1.6738*** (0.5685)	2.2907*** (0.7532)	1.4562* (0.7579)
R&D Credit (1 + cred)	3.2494*** (0.6970)	0.3407* (0.1974)	0.3580* (0.2145)	-0.0291 (0.2135)	0.0780 (0.2099)	0.3937* (0.2251)	1.0259*** (0.2789)	-0.4188 (0.3355)
Origin, Destination State FE	No	Yes	No	No	No	No	No	No
Origin*Destination Pair FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Origin Region*Year FE	No	No	No	Yes	No	No	No	No
Destination Region*Year FE	No	No	No	No	Yes	No	No	No
Origin & Destination Pair Region*Year FE	No	No	No	No	No	Yes	No	No
Origin State*Year FE	No	No	No	No	No	No	Yes	No
Destination State*Year FE	No	No	No	No	No	No	No	Yes

Notes: Each entry is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3). Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair. All regressions include year fixed effects, and have 15247 observations. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$



Table Appendix 7: Asymmetric Effects for Origin and Destination Taxes

	(1)	(2)
MTR, 99th Perc., Origin (1 - mtr)	-3.2053*** (0.9538)	-2.5043*** (0.8470)
MTR, 99th Perc., Dest. (1 - mtr)	1.4591** (0.6724)	2.2080*** (0.7055)
State CIT Rate, Origin (1 - cit)	-1.7581 (1.1944)	-2.0172* (1.1061)
State CIT Rate, Dest. (1 - cit)	1.4096* (0.7712)	1.2425 (0.7669)
State ITC, Origin (1 + itc)	-1.4770** (0.7464)	-1.2906* (0.7170)
State ITC, Dest. (1 + itc)	1.8780** (0.7325)	2.0809** (0.8361)
R&D Credit, Origin (1 + cred)	0.2497 (0.3330)	0.4044 (0.3298)
R&D Credit, Dest. (1 + cred)	0.9562*** (0.2955)	1.1790*** (0.3317)
Origin*Destination Pair FE	Yes	Yes
Origin & Destination Pair Region*Year FE	No	Yes

Notes: Each column is from a separate regression. Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair. All regressions include year fixed effects, and have 15247 observations. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table Appendix 8: Number of State Pairs with no flow, 1977-2006

	State Pairs	Average
1977	2178.00	210.49
1978	2150.00	208.92
1979	2171.00	203.82
1980	2235.00	155.39
1981	2219.00	151.98
1982	2235.00	147.18
1983	2170.00	155.68
1984	2194.00	165.06
1985	2153.00	182.20
1986	2150.00	188.59
1987	2110.00	219.88
1988	2140.00	199.25
1989	2097.00	220.45
1990	2045.00	244.92
1991	2066.00	262.49
1992	2016.00	295.65
1993	1994.00	330.78
1994	1934.00	392.65
1995	1941.00	324.35
1996	1972.00	355.69
1997	1918.00	410.31
1998	1928.00	376.80
1999	1900.00	428.69
2000	1877.00	407.41
2001	1894.00	431.61
2002	1889.00	444.51
2003	1932.00	394.06
2004	1948.00	376.88
2005	1948.00	353.82
2006	2037.00	280.54

Notes: First column shows the number of origin-destination-year cells that have zero scientist flows. Second column shows the average (across states) number of star scientists in that year.



Table Appendix 11:  
The Effect of Net-of-Tax Rates on Outmigration of Star Scientists: Estimates based on Alternative Definitions of Stars

Table Panel A: The Effect of Net-of-Tax Rates on Outmigration of Star Scientists: Baseline Models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATR, 99th Perc. (1 - atr)	2.1711** (1.0062)	1.9894** (0.8196)	1.9273** (0.8377)	2.2291*** (0.6351)	2.1966*** (0.8406)	1.9072** (0.8081)	0.8940 (0.7571)	4.2672*** (1.4613)
State CIT Rate (1 - cit)	-4.2411*** (1.0740)	2.8362*** (1.0136)	3.0721*** (1.0321)	2.5362*** (0.8615)	2.5736** (1.0528)	3.0951*** (0.9914)	1.0770 (0.8935)	6.0301*** (1.7128)
State ITC (1 + itc)	7.2826*** (1.7307)	1.7594*** (0.6023)	1.6399** (0.6863)	1.0752 (0.6763)	1.7093*** (0.6494)	1.5382** (0.7597)	1.9962** (0.9031)	1.5171 (1.0956)
R&D Credit (1 + cred)	3.6110*** (0.8491)	0.6105** (0.2916)	0.6817** (0.3257)	-0.1677 (0.2814)	0.4638 (0.3245)	0.6359* (0.3330)	0.8688** (0.3624)	0.0776 (0.5373)
Origin, Destination State FE	No	Yes	No	No	No	No	No	No
Origin*Destination Pair FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Origin Region*Year FE	No	No	No	Yes	No	No	No	No
Destination Region*Year FE	No	No	No	No	Yes	No	No	No
Origin & Destination Pair Region*Year FE	No	No	No	No	No	Yes	No	No
Origin State*Year FE	No	No	No	No	No	No	Yes	No
Destination State*Year FE	No	No	No	No	No	No	No	Yes

Notes: Each entry is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3). Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair. All regressions include year fixed effects, and have 7958 observations. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table Panel B: The Effect of Net-of-Tax Rates on Outmigration of Star Scientists: Baseline Models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATR, 99th Perc. (1 - atr)	2.8465*** (0.8214)	2.0929*** (0.6434)	1.8856*** (0.6743)	1.7649*** (0.5853)	2.0027*** (0.6655)	1.8575*** (0.6166)	1.1749 (0.7243)	2.2773** (1.0318)
State CIT Rate (1 - cit)	-4.0092*** (0.9159)	1.4978** (0.6767)	1.3203* (0.6826)	1.1113* (0.6123)	0.5670 (0.7052)	1.4325** (0.6583)	2.6323*** (0.7810)	0.3496 (1.0837)
State ITC (1 + itc)	4.6659*** (1.3812)	1.4266*** (0.4567)	1.3301** (0.5440)	1.3383** (0.5711)	1.3738*** (0.4926)	1.3277** (0.5634)	1.8880** (0.7959)	0.8679 (0.7029)
R&D Credit (1 + cred)	3.2193*** (0.6742)	0.3444* (0.1925)	0.3313 (0.2107)	0.0405 (0.2112)	0.0452 (0.2020)	0.3604* (0.2186)	1.1508*** (0.2898)	-0.5400* (0.3184)
Origin, Destination State FE	No	Yes	No	No	No	No	No	No
Origin*Destination Pair FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Origin Region*Year FE	No	No	No	Yes	No	No	No	No
Destination Region*Year FE	No	No	No	No	Yes	No	No	No
Origin & Destination Pair Region*Year FE	No	No	No	No	No	Yes	No	No
Origin State*Year FE	No	No	No	No	No	No	Yes	No
Destination State*Year FE	No	No	No	No	No	No	No	Yes

Notes: Each entry is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3). Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair. All regressions include year fixed effects, and have 17733 observations. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table Appendix 12: The Effect of Net-of-tax Rates on Outmigration of Star Scientists:  
 Estimates based on Definition of Stars that depend on Patent Citations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATR, 99th Perc. (1 - atr)	2.6186*** (0.8520)	2.2226*** (0.6415)	2.0234*** (0.6678)	1.9504*** (0.5584)	2.1599*** (0.6693)	2.0327*** (0.6276)	0.9862 (0.6339)	3.1391*** (1.0993)
State CIT Rate (1 - cit)	-3.7902*** (0.9604)	2.0080*** (0.7683)	1.8599** (0.7771)	1.2626* (0.6605)	1.2745 (0.8029)	1.9042*** (0.7347)	1.5694** (0.6929)	2.8082** (1.2983)
State ITC (1 + itc)	6.0377*** (1.5297)	1.9363*** (0.4946)	1.7948*** (0.5725)	1.6790*** (0.5839)	1.8019*** (0.5280)	1.7928*** (0.6110)	2.5816*** (0.7971)	1.5959* (0.8193)
R&D Credit (1 + cred)	3.5257*** (0.7304)	0.3718* (0.2133)	0.3856* (0.2305)	-0.1036 (0.2221)	0.0997 (0.2239)	0.4202* (0.2396)	0.9224*** (0.2926)	-0.3593 (0.3509)
Origin, Destination State FE	No	Yes	No	No	No	No	No	No
Origin*Destination Pair FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Origin Region*Year FE	No	No	No	Yes	No	No	No	No
Destination Region*Year FE	No	No	No	No	Yes	No	No	No
Origin & Destination Pair Region*Year FE	No	No	No	No	No	Yes	No	No
Origin State*Year FE	No	No	No	No	No	No	Yes	No
Destination State*Year FE	No	No	No	No	No	No	No	Yes

Notes: Each entry is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3). Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair. All regressions include year fixed effects, and have 14156 observations. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table Appendix 13: The Effect of Net-of-Tax Rates on Outmigration of Star Scientists: Estimates based on Alternative Levels of Personal Income

Table Panel A: Tax Variables Included Simultaneously, Star 95th Percentile

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATR, 95th Perc. (1 - atr)	5.0161*** (0.8339)	4.6004*** (0.7307)	4.5323*** (0.7617)	4.2035*** (0.6509)	4.7227*** (0.7591)	4.5654*** (0.6990)	1.9361*** (0.6672)	7.0223*** (1.1684)
State CIT Rate (1 - cit)	-4.6332*** (0.9332)	1.4632** (0.6705)	1.3548** (0.6678)	0.9650* (0.5768)	0.6218 (0.6939)	1.3997** (0.6200)	1.8682*** (0.6425)	1.1545 (1.0598)
State ITC (1 + itc)	6.4883*** (1.4511)	2.2648*** (0.4777)	2.1822*** (0.5582)	2.1447*** (0.5685)	2.2990*** (0.5173)	2.1927*** (0.5906)	2.4894*** (0.7702)	2.2902*** (0.7875)
R&D Credit (1 + cred)	3.2264*** (0.7130)	0.3004 (0.2007)	0.3309 (0.2184)	-0.0723 (0.2186)	0.0315 (0.2090)	0.3668 (0.2287)	1.0148*** (0.2908)	-0.4608 (0.3201)
Origin, Destination State FE	No	Yes	No	No	No	No	No	No
Origin*Destination Pair FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Origin Region*Year FE	No	No	No	Yes	No	No	No	No
Destination Region*Year FE	No	No	No	No	Yes	No	No	No
Origin & Destination Pair Region*Year FE	No	No	No	No	No	Yes	No	No
Origin State*Year FE	No	No	No	No	No	No	Yes	No
Destination State*Year FE	No	No	No	No	No	No	No	Yes

Standard errors in parentheses

Each entry is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3).

Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair.

All regressions include year fixed effects, and have 15247 observations.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table Panel B: Tax Variables Included Simultaneously, Star 95th Percentile

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATR, 99.9th Perc. (1 - atr)	1.5300* (0.8411)	1.3583*** (0.5216)	1.4612*** (0.5265)	1.4032*** (0.4644)	1.4974*** (0.5293)	1.4741*** (0.4886)	1.3182** (0.5336)	1.7562* (0.9214)
State CIT Rate (1 - cit)	-3.5540*** (0.9391)	2.1507*** (0.7150)	2.0109*** (0.7092)	1.5828*** (0.6131)	1.3424* (0.7439)	2.0550*** (0.6670)	2.1366*** (0.6726)	2.1374* (1.2279)
State ITC (1 + itc)	5.1192*** (1.4448)	1.7833*** (0.4617)	1.7049*** (0.5412)	1.6618*** (0.5651)	1.7513*** (0.4980)	1.7115*** (0.5732)	2.3368*** (0.7634)	1.5084** (0.7559)
R&D Credit (1 + cred)	3.2684*** (0.6946)	0.3586* (0.2023)	0.3764* (0.2194)	-0.0179 (0.2176)	0.0940 (0.2116)	0.4117* (0.2299)	1.0434*** (0.2874)	-0.4048 (0.3361)
Origin, Destination State FE	No	Yes	No	No	No	No	No	No
Origin*Destination Pair FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Origin Region*Year FE	No	No	No	Yes	No	No	No	No
Destination Region*Year FE	No	No	No	No	Yes	No	No	No
Origin & Destination Pair Region*Year FE	No	No	No	No	No	Yes	No	No
Origin State*Year FE	No	No	No	No	No	No	Yes	No
Destination State*Year FE	No	No	No	No	No	No	No	Yes

Standard errors in parentheses

Each entry is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3).

Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair.

All regressions include year fixed effects, and have 15247 observations.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table Appendix 14: Robustness Checks

	Log Odds Ratio (1)	Log Odds Ratio (2)	Log Odds Ratio (3)	Log Odds Ratio (4)	Log Odds Ratio (5)
ATR, 99th Perc. (1 - atr)	Baseline 1.8895*** (0.6160)	Weighted by Average Outflow 0.5175 (0.8604)	Pop. and UE controls 1.8551*** (0.5453)	Through 2006 2.1423*** (0.6298)	Disambiguation Data Set 2.0920*** (0.7776)
State CIT Rate (1 - cit)	1.9286*** (0.6615)	2.2219** (0.9245)	1.6670** (0.6868)	1.9955*** (0.7398)	1.8687** (0.8542)
State ITC (1 + itc)	1.7253*** (0.5825)	2.6359*** (0.7195)	1.5135*** (0.4208)	1.9729*** (0.5859)	1.1289 (0.7536)
R&D Credit (1 + cred)	0.3978* (0.2301)	1.0743*** (0.3349)	-0.0211 (0.1813)	0.3317 (0.2346)	0.3895 (0.3056)
No. Observations	15247	15247	14809	14809	7915
Origin*Destination Pair FE	Yes	Yes	Yes	Yes	Yes
Origin & Destination Pair					
Region*Year FE	Yes	Yes	Yes	Yes	Yes

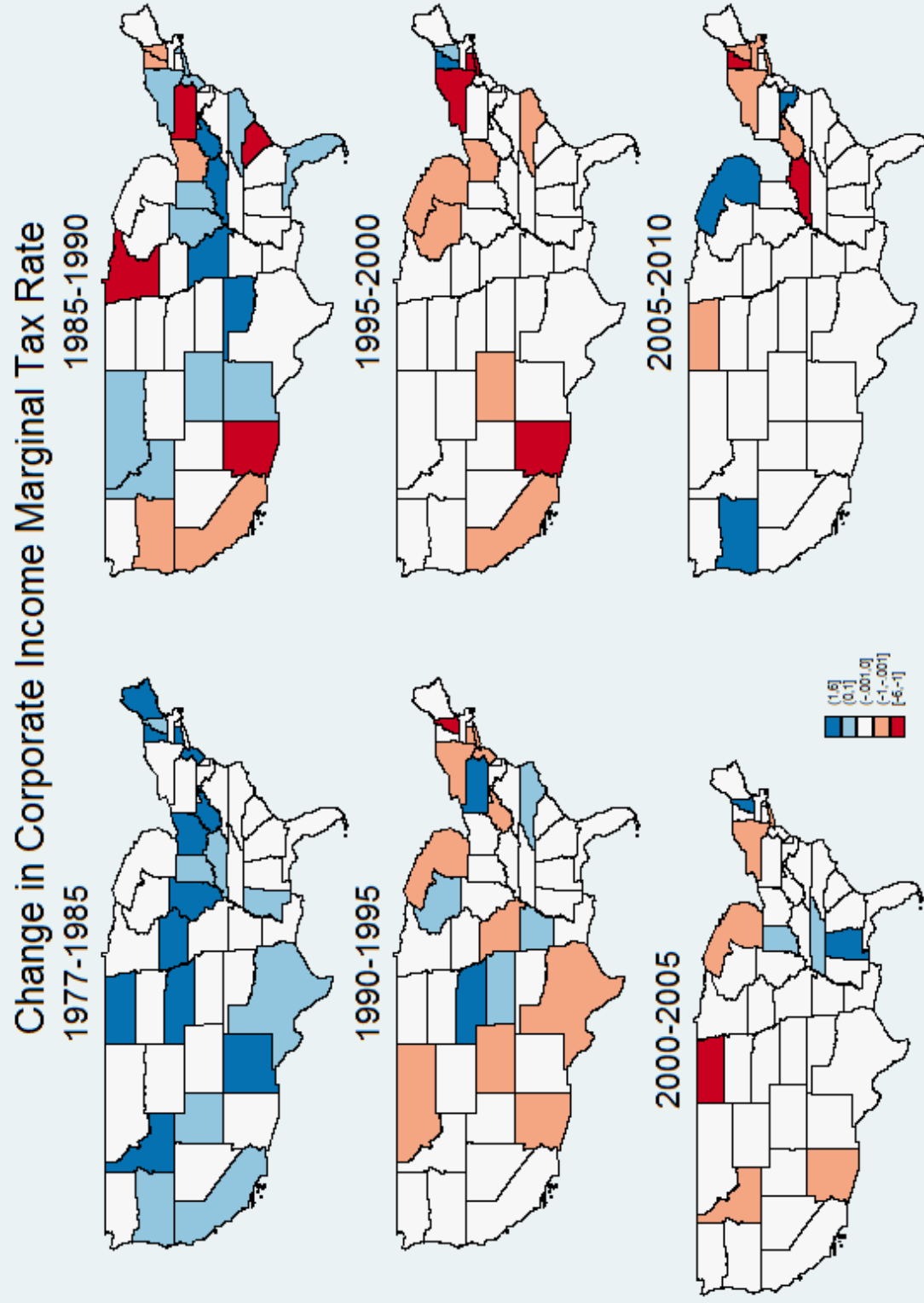
Notes: Each column is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3). Column 1 reproduces our baseline estimates from Table 2 Column 6. Column 2: Weighted by average outflow for origin-destination pair over sample. Column 3: Includes unemployment rate and population of origin and destination states as controls. Column 4: Sample excludes years after 2006. Column 5: Regression based on Li, et al. (2014) Disambiguation patent dataset. Standard errors in parentheses, with three-way clustering by origin-state\*year, destination-state\*year, and state-pair. All regressions include year fixed effects. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table Appendix 15: Scientist-Level Regression of Location Choice on Location Net-of-Tax Rate

	(1)	(2)
ATR, 99th Perc. (1 - atr)	0.3938*** (0.1431)	0.3938*** (0.1436)
State CIT Rate (1 - cit)	0.6990*** (0.1694)	0.6990*** (0.1699)
State ITC (1 + itc)	0.8597*** (0.0956)	0.8597*** (0.0959)
R&D Credit (1 + cred)	0.1410*** (0.0380)	0.1410*** (0.0381)
Origin*Destination Pair FE	Yes	Yes
Origin & Destination Pair		
Region*Year FE	No	Yes

Notes: Each column is from a separate regression. Coefficients are estimates of  $\eta$  or  $\eta'$  from equation (3). Driscoll-Kraay standard errors, with 1 lag, in parentheses. All regressions include year fixed effects, and have 22649151 observations. Coefficients and standard errors are multiplied by 100 for ease of display. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Appendix Figure 1

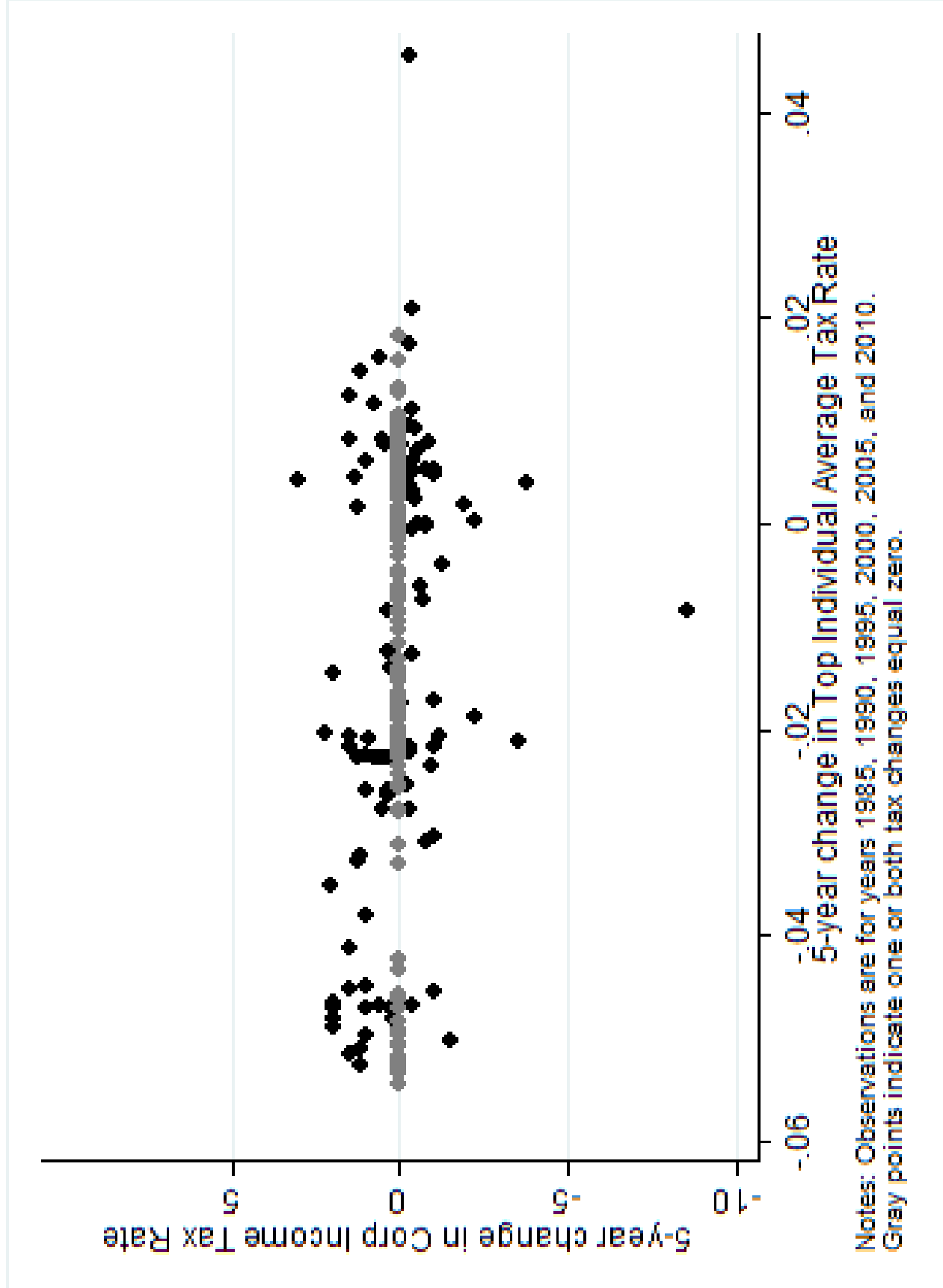


Notes: Categories are identical across maps. White indicates no change.



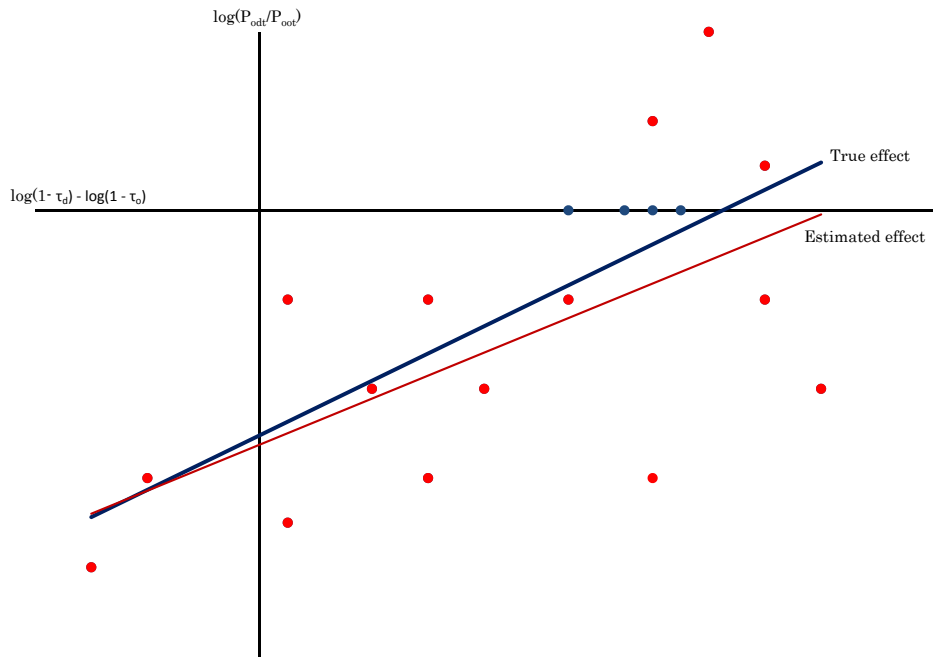


Appendix Figure 3

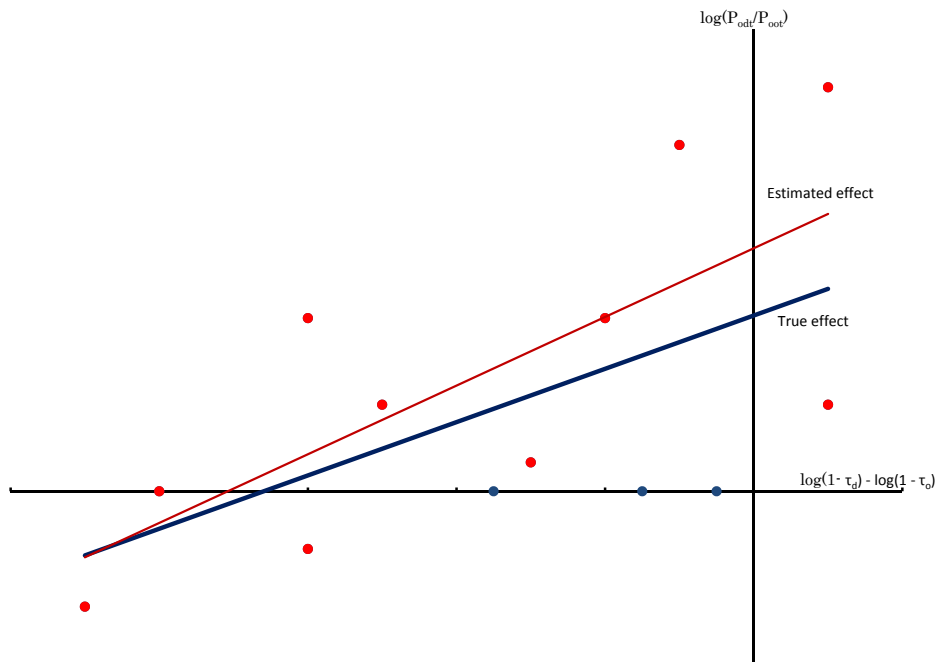


Appendix Figure 4

Example of Downward Bias



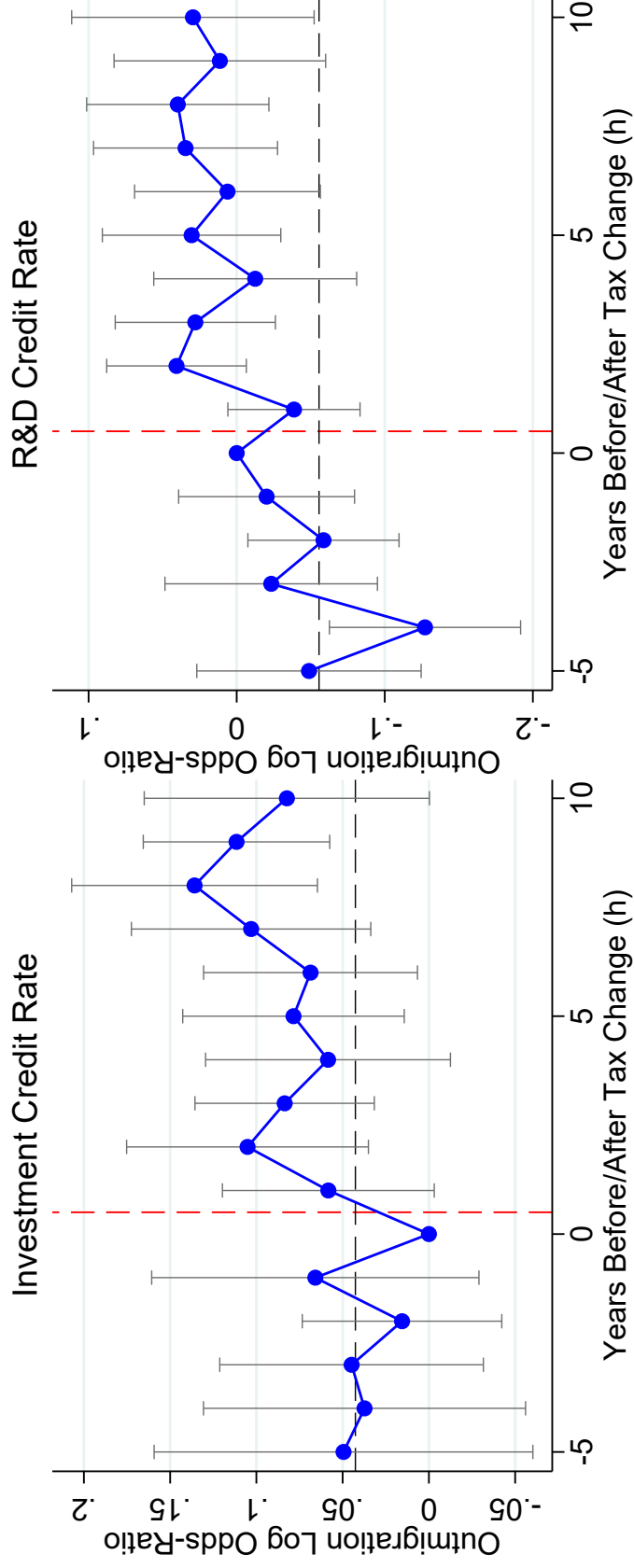
Example of Upward Bias



Notes: Each panel shows a hypothetical dataset. Each dot represents an origin-destination-year cell. The three blue points on the x-axis are not observed. The blue line has slope equal to the true effect of tax differentials on mobility. The blue line has the same slope in both panels. The slope of the red line is the estimated effect in the truncated sample.

Appendix Figure 5

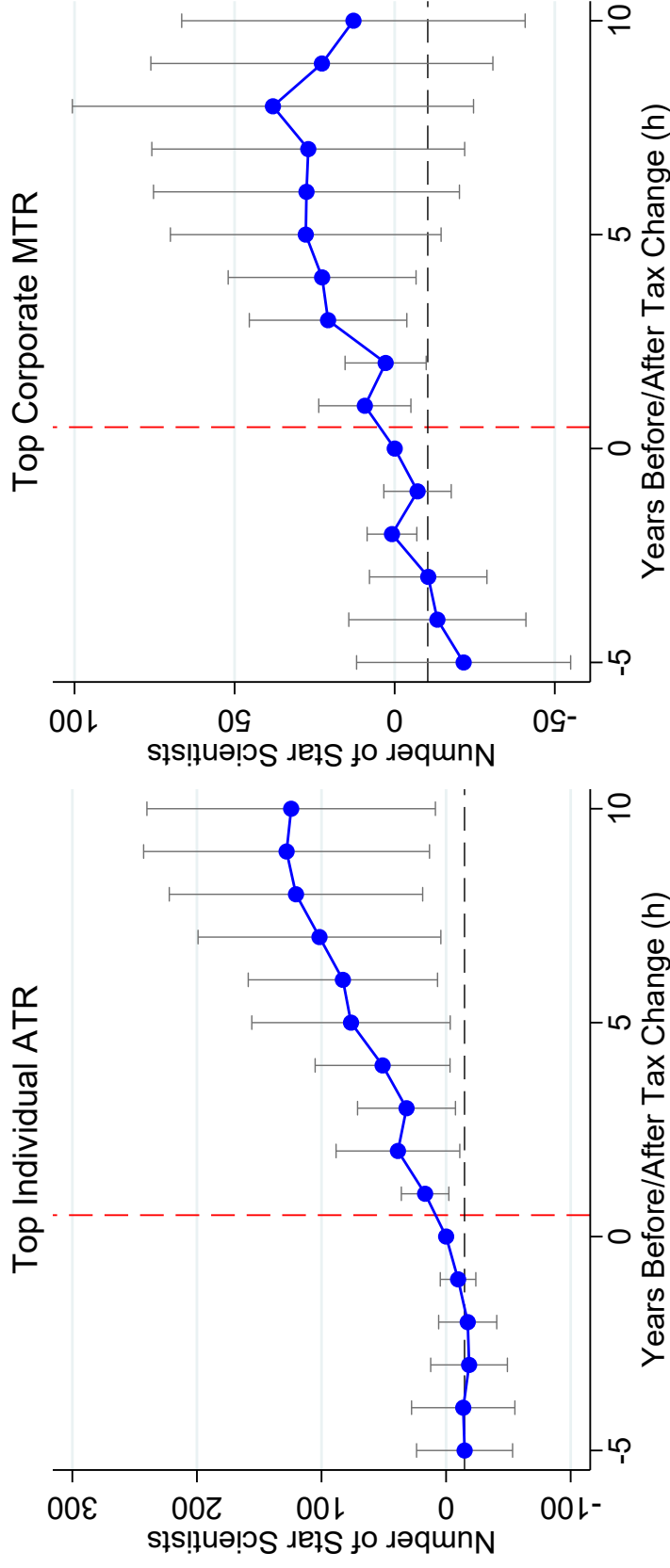
# Outmigration Before and After Tax Change Event



Notes: The dashed black line indicates the average coefficient over the pre-treatment period. We use a balanced panel from 5 years before event to 10 years after. The graph plots  $\beta^h$  from the regressions:  $\ln OR_{o,d,t+h} - \ln OR_{o,d,t} = \beta^h D_{o,d,t} + \varepsilon_{o,d,t}$ , where  $\ln OR$  is the outmigration log odds-ratio.  $D_{o,d,t}$  is an event indicator that takes the value 1 if the destination-origin differential in the net-of-tax rate increases between  $t$  and  $t+1$ , -1 if the differential decreases between  $t$  and  $t+1$ , and 0 if the differential does not change. Only permanent tax changes are included (defined as changes that are not reversed in the next 5 years).

Appendix Figure 6

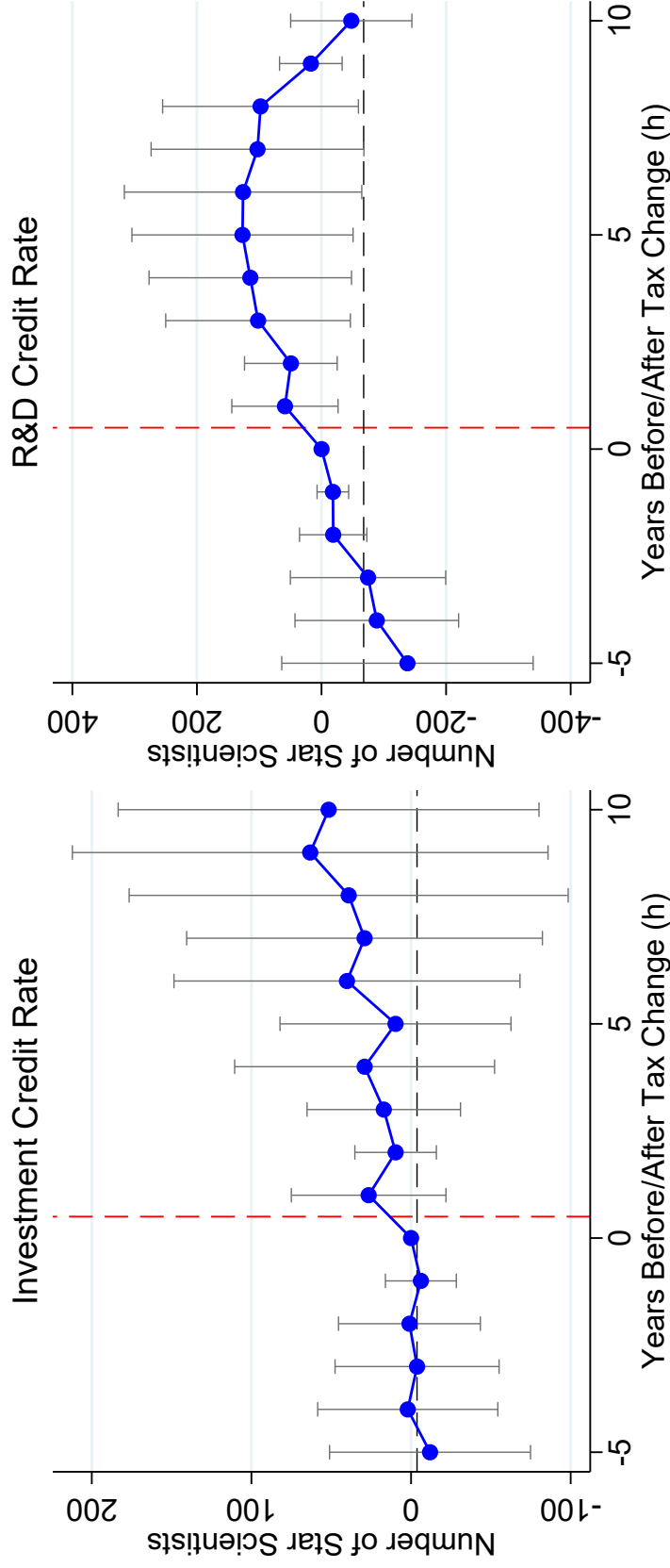
# Number of Stars Before and After Tax Change Event



Notes: The dashed black line indicates the average coefficient over the pre-treatment period. We use a balanced panel from 5 years before event to 10 years after. The graph plots  $\beta^h$  from the regressions:  $N_{s,t+h} - N_{s,t} = \beta^h D_{s,t} + \varepsilon_{s,t,h}$ , where  $N_{s,t}$  is the stock of star scientists in state  $s$  in year  $t$ .  $D_{s,t}$  is an event indicator that equals 1 (-1) if net-of-tax rate in  $s$  increases (decreases) between  $t$  and  $t+1$  and equals 0 if net-of-tax rate is unchanged. Only permanent tax changes are included (defined as changes that are not reversed in the next 5 years).

Appendix Figure 7

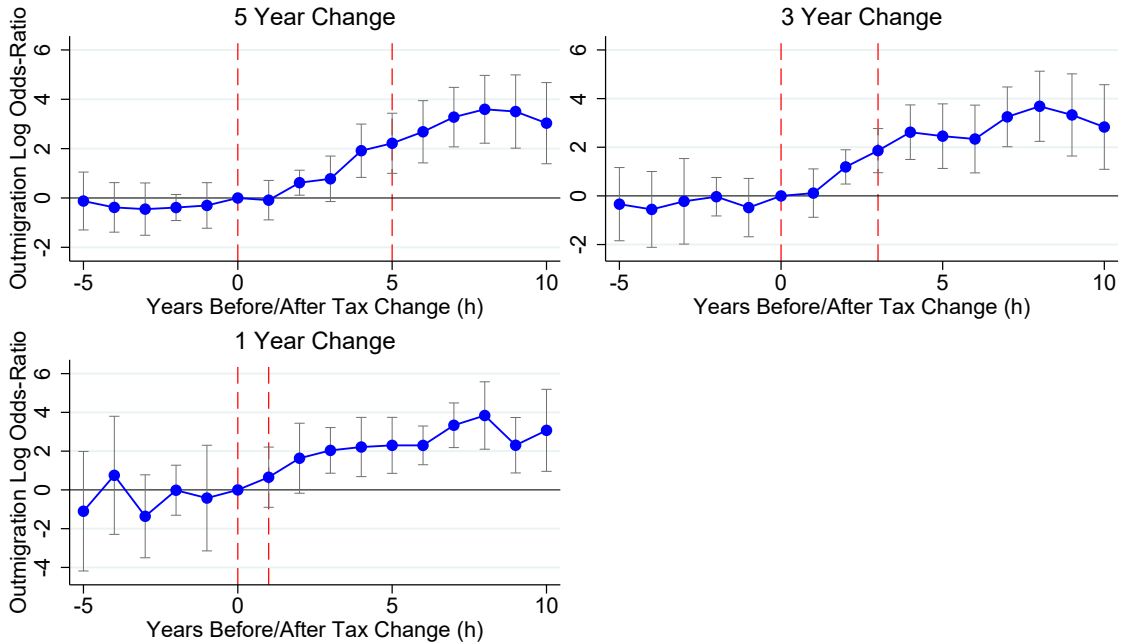
# Number of Stars Before and After Tax Change Event



Notes: The dashed black line indicates the average coefficient over the pre-treatment period. We use a balanced panel from 5 years before event to 10 years after. The graph plots  $\beta^h$  from the regressions:  $N_{s,t+h} - N_{s,t} = \beta^h D_{s,t} + \varepsilon_{s,t,h}$ , where  $N_{s,t}$  is the stock of star scientists in state  $s$  in year  $t$ .  $D_{s,t}$  is an event indicator that equals 1 (-1) if net-of-tax rate in  $s$  increases (decreases) between  $t$  and  $t+1$  and equals 0 if net-of-tax rate is unchanged. Only permanent tax changes are included (defined as changes that are not reversed in the next 5 years).

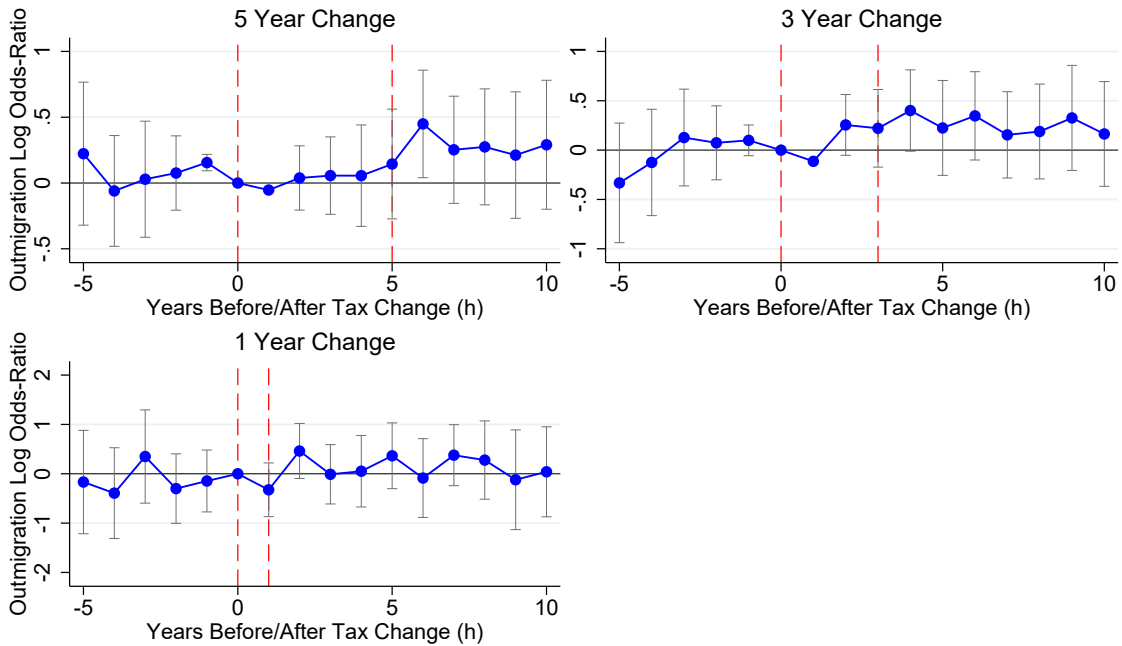
Appendix Figure 8. Impulse Response Functions

### Outmigration Response to Investment Credit Change



Notes: All regressions use the same sample, a balanced panel from 5 years before event to 10 years after. Graph plots  $\beta^h$  from regression:  $\ln P_{odt+h} - \ln P_{odt} = \beta^h (\nabla n_{odt+x} - \nabla n_{odt}) + F_{odt} + \varepsilon_{odt}$ , where  $\ln P$  is the log odds-ratio of outmigration and  $\nabla n_{odt}$  is the destination-origin differential in the net-of-tax rate in year t, and x is 1, 3, or 5.  $F_{odt}$  is a year fixed effect specific to the region(o)\*region(d) pair.

### Outmigration Response to R&D Credit Change



Notes: All regressions use the same sample, a balanced panel from 5 years before event to 10 years after. Graph plots  $\beta^h$  from regression:  $\ln P_{odt+h} - \ln P_{odt} = \beta^h (\nabla n_{odt+x} - \nabla n_{odt}) + F_{odt} + \varepsilon_{odt}$ , where  $\ln P$  is the log odds-ratio of outmigration and  $\nabla n_{odt}$  is the destination-origin differential in the net-of-tax rate in year t, and x is 1, 3, or 5.  $F_{odt}$  is a year fixed effect specific to the region(o)\*region(d) pair.