# Online Appendix to "State Taxes and Spatial Misallocation"

## A Appendix to Section 3

Figure A.1: Dispersion in State and Local Tax Rates in 2007

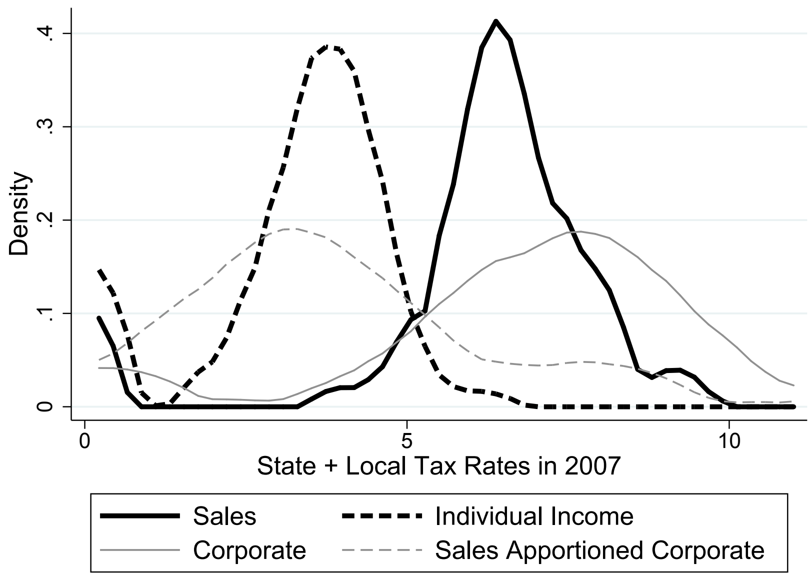



Table A.1: Federal Tax Rates in 2007

| Type | Federal Tax Rate |
|---|---|
| Income Tax $t^{y}_{fed}$ | 11.7 |
| Corporate Tax $t^{corp}_{fed}$ | 18 |
| Payroll Tax $t^{w}_{fed}$ | 7.3 |

Notes: This table shows federal tax rates in 2007 for individual income, corporate, and payroll taxes. The income tax rate is the average effective federal tax rate from NBER's TAXSIM across all states in 2007. The TAXSIM data we use provides the effective federal tax rate on individual income after accounting for deductions. The corporate tax rate is the average effective corporate tax rate: we divide total tax liability (including tax credits) by net business income less deficit, using data from IRS Statistics of Income on corporation income tax returns. Finally, for payroll tax rates, we use data from the Congressional Budget Office on federal tax rates for all households in 2007. This payroll rate is similar to the employer portion of the sum of Old-Age, Survivors, and Disabilty Insurance and Medicare's Hospital Insurance Program.

Table A.2: State Tax Rates in 2007

| State | $t_n^y$ | $t_n^c$ | $t_n^{corp}$ | $t_n^x$ |
|---|---|---|---|---|
| AL | 3.1 | 4 | 6.5 | 2.2 |
| AZ | 2.2 | 5.6 | 7 | 4.2 |
| AR | 3.7 | 6 | 6.5 | 3.2 |
| CA | 4 | 7.2 | 8.8 | 4.4 |
| CO | 3.3 | 2.9 | 4.6 | 1.5 |
| CT | 4 | 6 | 7.5 | 3.8 |
| DE | 3.5 | 0 | 8.7 | 2.9 |
| FL | 0 | 6 | 5.5 | 2.7 |
| GA | 4 | 4 | 6 | 5.4 |
| HI | 4.5 | 4 | 6.4 | 2.1 |
| ID | 4.5 | 6 | 7.6 | 3.8 |
| IL | 2.3 | 6.3 | 4.8 | 4.8 |
| IN | 3.1 | 6 | 8.5 | 5.1 |
| IA | 4.2 | 5 | 12 | 12 |
| KS | 4.1 | 5.3 | 7.3 | 2.5 |
| KY | 4.1 | 6 | 7 | 3.5 |
| LA | 3.1 | 4 | 8 | 8 |
| ME | 4.6 | 5 | 8.9 | 8.9 |
| MD | 3.5 | 6 | 7 | 3.5 |
| MA | 4.5 | 5 | 9.5 | 4.7 |
| MI | 3.1 | 6 | 1.9 | 1.8 |
| MN | 4.8 | 6.5 | 9.8 | 7.6 |
| MS | 2.8 | 7 | 5 | 1.7 |
| MO | 3.5 | 4.2 | 6.3 | 2.1 |
| MT | 3.7 | 0 | 6.8 | 2.3 |
| NE | 3.9 | 5.5 | 7.8 | 7.8 |
| NV | 0 | 6.5 | 0 | 0 |
| NH | .2 | 0 | 8.5 | 4.3 |
| NJ | 4.2 | 7 | 9 | 4.5 |
| NM | 2.9 | 5 | 7.6 | 2.5 |
| NY | 4.8 | 4 | 7.5 | 7.5 |
| NC | 5 | 4.3 | 6.9 | 3.4 |
| ND | 2.1 | 5 | 7 | 2.3 |
| OH | 3.5 | 5.5 | 8.5 | 5.1 |
| OK | 3.5 | 4.5 | 6 | 2 |
| OR | 6 | 0 | 6.6 | 6.6 |
| PA | 2.9 | 6 | 10 | 7 |
| RI | 3.6 | 7 | 9 | 3 |
| SC | 3.6 | 6 | 5 | 5 |
| SD | 0 | 4 | 0 | 0 |
| TN | .3 | 7 | 6.5 | 3.2 |
| TX | 0 | 6.3 | 0 | 0 |
| UT | 4 | 4.7 | 5 | 2.5 |
| VT | 3.4 | 6 | 8.5 | 4.3 |
| VA | 4.1 | 5 | 6 | 3 |
| WA | 0 | 6.5 | 0 | 0 |
| WV | 4.2 | 6 | 8.7 | 4.4 |
| WI | 4.5 | 5 | 7.9 | 6.3 |
| WY | 0 | 4 | 0 | 0 |

NOTES: This table shows state tax rates in 2007 for individual income ($t_n^y$), general sales ($t_n^c$), corporate ($t_n^{corp}$), and sales-apportioned corporate ($t_n^x$) taxes, which is the product of the statutory corporate tax rate and the state's sales apportionment weight. See Section 3.1 for details.

Table A.3: State Income Tax Parameters and Effective Tax Rates in 2007

| State | $a_{n,state}$ | $b_{n,state}$ | State tax rates if AGI is | | | | Overall tax rates if AGI is | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 25K | 50K | 100K | 200K | 25K | 50K | 100K | 200K |
| AL | 1.025 | 0.005 | 2.0 | 2.4 | 3.1 | 3.4 | 14.8 | 18.1 | 22.9 | 25.2 |
| AK | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| AZ | 1.078 | 0.008 | 0.2 | 1.0 | 2.2 | 2.7 | 13.4 | 17.0 | 22.2 | 24.7 |
| AR | 1.092 | 0.011 | 1.2 | 2.1 | 3.6 | 4.3 | 14.1 | 17.9 | 23.3 | 25.9 |
| CA | 1.102 | 0.011 | 0.3 | 1.3 | 2.7 | 3.5 | 13.4 | 17.2 | 22.7 | 25.3 |
| CO | 1.066 | 0.008 | 1.3 | 2.0 | 3.2 | 3.7 | 14.2 | 17.8 | 23.0 | 25.5 |
| CT | 1.087 | 0.010 | 0.9 | 1.8 | 3.2 | 3.9 | 13.9 | 17.7 | 23.0 | 25.6 |
| DE | 1.067 | 0.008 | 1.0 | 1.7 | 2.9 | 3.4 | 14.0 | 17.6 | 22.8 | 25.2 |
| FL | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| GA | 1.132 | 0.014 | 0.8 | 2.1 | 4.0 | 5.0 | 13.8 | 17.9 | 23.7 | 26.4 |
| HI | 1.136 | 0.015 | 1.2 | 2.5 | 4.5 | 5.5 | 14.1 | 18.2 | 24.1 | 26.8 |
| ID | 1.166 | 0.017 | 0.5 | 2.1 | 4.4 | 5.5 | 13.6 | 17.9 | 24.0 | 26.8 |
| IL | 1.019 | 0.004 | 1.4 | 1.8 | 2.3 | 2.5 | 14.3 | 17.6 | 22.3 | 24.6 |
| IN | 1.019 | 0.004 | 2.2 | 2.6 | 3.2 | 3.5 | 15.0 | 18.3 | 23.1 | 25.3 |
| IA | 1.122 | 0.014 | 1.2 | 2.5 | 4.3 | 5.2 | 14.2 | 18.2 | 23.9 | 26.6 |
| KS | 1.066 | 0.009 | 1.6 | 2.4 | 3.6 | 4.2 | 14.5 | 18.1 | 23.3 | 25.8 |
| KY | 1.070 | 0.009 | 1.9 | 2.7 | 4.0 | 4.6 | 14.7 | 18.4 | 23.6 | 26.1 |
| LA | 1.082 | 0.010 | 1.0 | 1.9 | 3.2 | 3.8 | 14.0 | 17.7 | 23.0 | 25.5 |
| ME | 1.131 | 0.015 | 1.2 | 2.5 | 4.5 | 5.5 | 14.1 | 18.2 | 24.0 | 26.8 |
| MD | 1.055 | 0.007 | 1.5 | 2.2 | 3.2 | 3.7 | 14.4 | 17.9 | 23.0 | 25.4 |
| MA | 1.055 | 0.008 | 2.4 | 3.1 | 4.2 | 4.8 | 15.1 | 18.7 | 23.8 | 26.2 |
| MI | 1.049 | 0.007 | 1.5 | 2.1 | 3.1 | 3.5 | 14.4 | 17.9 | 22.9 | 25.3 |
| MN | 1.108 | 0.013 | 1.4 | 2.5 | 4.2 | 5.1 | 14.3 | 18.2 | 23.8 | 26.5 |
| MS | 1.010 | 0.003 | 1.4 | 1.6 | 1.9 | 2.1 | 14.3 | 17.5 | 22.1 | 24.3 |
| MO | 1.065 | 0.008 | 1.3 | 2.0 | 3.1 | 3.7 | 14.2 | 17.8 | 23.0 | 25.4 |
| MT | 1.093 | 0.011 | 1.1 | 2.1 | 3.6 | 4.3 | 14.1 | 17.9 | 23.3 | 25.9 |
| NE | 1.109 | 0.012 | 0.8 | 1.9 | 3.6 | 4.4 | 13.8 | 17.7 | 23.3 | 25.9 |
| NV | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| NH | 1.000 | 0.000 | 0.1 | 0.1 | 0.1 | 0.1 | 13.2 | 16.3 | 20.7 | 22.8 |
| NJ | 1.054 | 0.007 | 0.8 | 1.4 | 2.3 | 2.8 | 13.8 | 17.3 | 22.4 | 24.8 |
| NM | 1.183 | 0.017 | -0.8 | 0.8 | 3.1 | 4.3 | 12.5 | 16.8 | 23.0 | 25.9 |
| NY | 1.099 | 0.012 | 1.3 | 2.4 | 4.0 | 4.7 | 14.3 | 18.1 | 23.6 | 26.2 |
| NC | 1.055 | 0.009 | 2.5 | 3.2 | 4.4 | 5.0 | 15.2 | 18.8 | 23.9 | 26.4 |
| ND | 1.052 | 0.006 | 0.5 | 1.0 | 1.8 | 2.2 | 13.5 | 17.0 | 22.0 | 24.4 |
| OH | 1.061 | 0.008 | 1.2 | 1.9 | 2.9 | 3.4 | 14.1 | 17.7 | 22.8 | 25.2 |
| OK | 1.146 | 0.016 | 0.7 | 2.1 | 4.2 | 5.2 | 13.7 | 17.9 | 23.8 | 26.6 |
| OR | 1.107 | 0.014 | 2.7 | 4.0 | 5.8 | 6.7 | 15.4 | 19.4 | 25.0 | 27.7 |
| PA | 1.046 | 0.007 | 1.5 | 2.1 | 3.0 | 3.5 | 14.4 | 17.9 | 22.9 | 25.3 |
| RI | 1.095 | 0.011 | 0.8 | 1.7 | 3.2 | 3.9 | 13.8 | 17.6 | 23.0 | 25.6 |
| SC | 1.071 | 0.009 | 1.1 | 1.9 | 3.0 | 3.6 | 14.1 | 17.7 | 22.9 | 25.4 |
| SD | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| TN | 1.001 | 0.000 | 0.1 | 0.1 | 0.2 | 0.2 | 13.2 | 16.3 | 20.7 | 22.8 |
| TX | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| UT | 1.087 | 0.011 | 1.4 | 2.3 | 3.8 | 4.5 | 14.3 | 18.1 | 23.5 | 26.0 |
| VT | 1.177 | 0.017 | -0.5 | 1.1 | 3.4 | 4.6 | 12.8 | 17.1 | 23.2 | 26.1 |
| VA | 1.076 | 0.010 | 1.6 | 2.4 | 3.7 | 4.4 | 14.4 | 18.1 | 23.4 | 25.9 |
| WA | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| WV | 1.062 | 0.009 | 1.9 | 2.7 | 3.9 | 4.4 | 14.8 | 18.4 | 23.5 | 26.0 |
| WI | 1.086 | 0.011 | 1.8 | 2.8 | 4.2 | 4.9 | 14.6 | 18.4 | 23.8 | 26.4 |
| WY | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |

NOTES: This table shows state income tax parameters in 2007 as well as effective tax rates for different levels of Adjusted Gross Income (AGI). Tax rates reported in columns 4-7 are state-only, while tax rates in columns 8-11 combine federal and state taxation. Federal taxation includes individual income taxes and the employee portion of payroll (FICA) taxes.

## A.1 Supplemental Stylized Facts on State Taxes

Figure A.2: Supplemental Stylized Facts on State Taxes


Notes: Panel (a) plots state government direct expenditure against model-based tax revenue in 2007. Data are drawn from Census Government Finances. Panel (b) plots the statutory state corporate keep rate, as measured by $1 - t_n^{corp}$, against the combined federal and state effective individual income keep rate, which is estimated using NBER's tax simulator TAXSIM. For each state, we compute average federal and state tax liabilities and divide their sum by average Adjusted Gross Income (AGI) in that state. Then, we account for the impact of sales taxes on individuals' purchasing power by dividing the raw keep rate by $1 + t_n^c$. Panel (c) shows the correlation between the combined federal and state individual income keep rate and the statutory state sales keep rate, as measured by $1 - t_n^c$. We estimate the former using NBER's tax simulator TAXSIM. For each state, we compute average federal and state tax liabilities and divide their sum by average Adjusted Gross Income (AGI) in that state. Panel (d) plots the statutory state corporate keep rate against the statutory state sales keep rate. In panels (b), (c), and (d), the vertical and horizontal grey lines denote population-weighted averages of the variables on the x- and y-axis, respectively. Observations are weighted by state population.

## A.2 Bilateral Trade Shares and State Corporate Taxation

According to our model, using (A.3), (A.4), and (A.5) and aggregating across firms, inter-state trade flows take the form:

$$\ln X_{ni} = \beta_1 \ln \frac{\sigma}{\sigma - \tilde{t}_{ni}} + \beta_2 \ln \tau_{ni} + \psi_i + \psi_n + \varepsilon_{in},$$

where $\psi_i$ and $\psi_n$ are respectively origin and destination fixed effects. As shown in (A.3), $\tilde{t}_{ni}$ is a function of the matrix of trade flows and corporate taxes and therefore we instrument for this term using corporate taxes in the destination only.

Table A.4: Bilateral Trade Shares and Trade-Dispersion Cost

| | (1) OLS | (2) OLS | (3) OLS | (4) OLS | (5) 2SLS | (6) 2SLS | (7) 2SLS | (8) 2SLS |
|---|---|---|---|---|---|---|---|---|
| $\ln \frac{\sigma}{\sigma-\tilde{t}}$ | -4.265* | -3.414 | -2.250 | -2.948** | -9.513*** | -3.993 | -2.631 | -2.590* |
| | (2.204) | (2.111) | (2.154) | (1.289) | (2.660) | (2.448) | (2.491) | (1.390) |
| Observations | 10,512 | 10,512 | 10,512 | 10,272 | 10,512 | 10,512 | 10,512 | 10,272 |
| R-squared | 0.457 | 0.474 | 0.474 | 0.826 | 0.456 | 0.474 | 0.474 | 0.826 |
| Year FE | No | Yes | Yes | Yes | No | Yes | Yes | Yes |
| Dest GDP Control | No | No | Yes | Yes | No | No | Yes | Yes |
| Distance Control | No | No | No | Yes | No | No | No | Yes |

NOTES: The panel consists of the 48 contiguous states in 1993, 1997, 2002, 2007, and 2012. Each observation is an origin-destination-year triplet. In all specifications, the dependent variable is log bilateral trade share, which is defined as $s_{in} = \frac{x_{in}}{\sum_i x_{in}}$, where $x_{in}$ denotes sales from state $n$ to state $i$. All models allow for origin and destination state fixed effects. Observations are weighted by destination state population. Columns 1-4 show the association between $\ln \frac{\sigma}{\sigma-\tilde{t}}$ and bilateral trade share, allowing for year fixed effects (Column 2), and controlling for destination state GDP (Column 3) and distance between state pairs (Column 4). In Column 5, $\ln \frac{\sigma}{\sigma-\tilde{t}}$ is instrumented with destination $t^x$. In Column 6, this specification is augmented with year fixed effects. Columns 7 and 8 also control for destination state GDP and distance between state pairs, respectively. Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

# B Appendix to Section 4

## B.1 Firm Maximization

We characterize here the problem in (20) for a firm $j$ located in $i$ whose productivity is $z$. When a firm $j$ located in state $i$ sets its price $p^j_{ni}$ in state $n$, the quantity exported to state $n$ is $q^j_{ni} = Q_n(p^j_{ni}/P_n)^{-\sigma}$. The first-order condition of profits (20) with respect the quantity sold to $n$ is:

$$\frac{\partial \pi^j_i}{\partial q^j_{ni}} = (1 - \bar{t}^j_i)\frac{\partial \tilde{\pi}^j_i}{\partial q^j_{ni}} - \frac{\partial \bar{t}^j_i}{\partial q^j_{ni}}\tilde{\pi}^j_i = 0, \tag{A.1}$$

where $\tilde{\pi}^j_i \equiv \sum_{n=1}^N x^j_{ni} - (\tau_{ni}c_i/z)q^j_{ni}$ are pre-tax profits, and where:

$$\frac{\partial \tilde{\pi}^j_i}{\partial q^j_{ni}} = \frac{\sigma-1}{\sigma} E_n^{1/\sigma} P_n^{1-1/\sigma} \left(q^j_{ni}\right)^{-1/\sigma} - c_i \frac{\tau_{ni}}{z},$$

$$\frac{\partial \bar{t}^j_i}{\partial q^j_{ni}} = \frac{\sigma-1}{\sigma} \left( t^x_n - \sum_{n'} t^x_{n'} s^j_{n'i} \right) \frac{p^j_{ni}}{x^j_i}.$$

Combining the last two expressions with (A.1) yields:

$$p_{ni}^j = \frac{1}{1-\tilde{t}_{ni}^j\left(\tilde{\pi}_i^j/x_i^j\right)}\frac{\sigma}{\sigma-1}\frac{\tau_{ni}}{z}c_i, \tag{A.2}$$

where

$$\tilde{t}_{ni}^j \equiv \frac{t_n^x - \sum_{n'} t_{n'}^x s_{n'i}^j}{1-\bar{t}_i}. \tag{A.3}$$

Expressing pre-tax profits as

$$\tilde{\pi}_i^j \equiv \sum_{n=1}^N x_{ni}^j \left(1 - \frac{\tau_{ni}}{z}\frac{c_i}{p_{ni}^j}\right),$$

introducing this expression in (A.2) and using that $\sum_i s_{ni}^j \tilde{t}_{ni}^j = 0$ yields $\tilde{\pi}_i^j = x_i^j/\sigma$. This implies that

$$p_{ni}^j = \frac{\sigma}{\sigma - \tilde{t}_{ni}^j}\frac{\sigma}{\sigma-1}\frac{\tau_{ni}}{z}. \tag{A.4}$$

Finally, note that export shares are independent of productivity, $z_i^j$:

$$s_{ni}^j = \frac{E_n\left(p_{ni}^j\right)^{1-\sigma}}{\sum_{n'=1}^N E_{n'}\left(p_{n'i}^j\right)^{1-\sigma}} = \frac{E_n\left(\frac{\sigma-\tilde{t}_{ni}^j}{\tau_{ni}}\right)^{\sigma-1}}{\sum_{n'=1}^N E_{n'}\left(\frac{\sigma-\tilde{t}_{n'i}^j}{\tau_{n'i}}\right)^{1-\sigma}}. \tag{A.5}$$

Equations (A.3) and (A.5) for $n = 1,..,N$ define a system for $\{\tilde{t}_{ni}^j\}$ and $\{s_{ni}^j\}$ whose solution is independent of the firm's productivity $z$. Therefore, $\tilde{t}_{ni}^j = \tilde{t}_{ni}$ and $s_{ni}^j = s_{ni}$ for all firms $j$ located in state $i$.

## B.2 Additional State-Level Variables

In this section, we let $\chi^{FE}$ be an indicator variable that equals 1 when we assume free entry of homogeneous firms and zero when we assume free mobility of heterogeneous firms.

**Factor Payments** From the Cobb-Douglas technologies and CES demand, in addition to the free-entry condition when $\chi^{FE} = 1$, it follows that payments to intermediate inputs, labor and fixed factors in state $i$ can be expressed as fractions of sales $X_i$:

$$P_i I_i = \left(1 + \chi^{FE}\frac{1-\bar{t}_i}{\sigma-1}\right)(1-\gamma_i)\frac{\sigma-1}{\sigma}X_i, \tag{A.6}$$

$$w_i L_i^E = \left(1 + \chi^{FE}\frac{1-\bar{t}_i}{\sigma-1}\right)(1-\beta_i)\gamma_i\frac{\sigma-1}{\sigma}X_i, \tag{A.7}$$

$$r_i H_i = \left(1 + \chi^{FE}\frac{1-\bar{t}_i}{\sigma-1}\right)\beta_i\gamma_i\frac{\sigma-1}{\sigma}X_i. \tag{A.8}$$

In each of these expressions, the term multiplied by $\chi^{FE}$ reflects the resources devoted to pay for entry costs. In the second equation, $L_i^E$ are the total efficiency units of labor demanded in state $i$, in equilibrium these efficiency units equal $L_i h_i E_i[z]$, where $h_i$ are the hours worked by a worker with productivity $z$ in state $i$ in (13).

Aggregate pre-tax profits $\tilde{\Pi}_i$ are:

$$\tilde{\Pi}_i = \frac{X_i}{\sigma}, \tag{A.9}$$

After-tax profits, gross of entry costs when $\chi^{FE} = 1$, are:

$$\Pi_i = \left(1 - \bar{t}_n\right)\frac{X_i}{\sigma}. \tag{A.10}$$

**Expenditure and Sales Shares** The share of aggregate expenditures in state $n$ on goods produced in state $i$ is

$$\lambda_{ni} = M_i^{1+\frac{1-\chi^{FE}}{\varepsilon_F}(1-\sigma)} \left( \frac{\sigma}{\sigma - \tilde{t}_{ni}} \frac{\sigma}{\sigma - 1} \frac{\tau_{ni} c_i}{z_i^0} \frac{1}{P_n} \right)^{1-\sigma}. \tag{A.11}$$

Under free entry ($\chi^{FE} = 1$), the congestion effect from entry on productivity described in Section 4.4 is absent. We construct the sales shares $s_{ni}$, which are necessary to compute the corporate tax rate $\bar{t}_i$ in (22) and the pricing distortion $\tilde{t}_{ni}$ in (24), using the identity $s_{ni} = \lambda_{ni} P_n Q_n / X_i$, where $P_n Q_n$ is the aggregate expenditure on final goods in state $n$.

**Real GDP** Adding up (A.7), (A.8), and (A.9), in the case with ex-ante heterogeneous firms, real GDP in state $n$ is

$$\frac{GDP_n}{P_n} = \frac{1 + \gamma_n (\sigma - 1) \left(1 - (1 - \beta_n) t^w_{fed} / \left(1 + t^w_{fed}\right)\right)}{\sigma} \frac{X_n}{P_n}. \tag{A.12}$$

Aggregate real GDP is defined as $GDP^{real} = \sum_n (GDP_n / P_n)$.

**Consumption** The aggregate personal consumption expenditure in state $n$ is $P_n C_n = P_n C_n^W + P_n C_n^K$, where $C_n^W$ is the aggregate real consumption of workers and $C_n^K$ is the consumption of capital owners. Taking into account the taxes paid to each level of government, these aggregates are:

$$P_n C_n^W = \mathbb{E}_n \left[ \frac{1 - T_n^y (w_n h_n z)}{1 + t_n^c} w_n h_n z \right] L_n \tag{A.13}$$

$$P_n C_n^K = \frac{\left(1 - \chi^{FE}\right) \tilde{\Pi} + R - T^{corp} - \left(\overline{t_n^y} + \overline{t_{n,fed}^y} \left(1 - \overline{t_n^y}\right)\right) \left(\left(1 - \chi^{FE}\right) \Pi + R\right)}{1 + t_n^c} \omega_n, \tag{A.14}$$

where $\overline{t_n^y}$ and $\overline{t_{n,fed}^y}$ are the top average state and federal personal income tax rates, $\Pi = \sum_i \Pi_i$, $\tilde{\Pi} = \sum_i \tilde{\Pi}_i$ and $R = \sum_i r_i H_i$ are national after-tax profits, pre-tax profits and returns to land and structures, respectively, and $T^{corp}$ are the national corporate tax payments.

**State Tax Revenue By Type of Tax** State government revenue from corporate, sales, and income taxes, is, respectively,

$$R_n^{corp} = t_n^x \sum_{n'} s_{nn'} \tilde{\Pi}_{n'} + t_n^l \tilde{\Pi}_n, \tag{A.15}$$

$$R_n^y = \mathbb{E} \left[ t_n^y (w_n h_n z) w_n h_n z \right] L_n + \overline{t_n^y} \omega_n \left( \left(1 - \chi^{FE}\right) \Pi + R \right), \tag{A.16}$$

$$R_n^c = t_n^c P_n C_n. \tag{A.17}$$

The base for corporate tax revenues are the pre-tax profits from every state, defined in (A.9), adjusted by the proper apportionment weights. Equation (A.16) shows that the base for state income taxes is the income of both workers and capital-owners who reside in $n$ net of federal income taxes. Income tax revenue from workers results from aggregating tax payments over the distribution of individual productivity. Capital owners are at the highest rate, $\overline{t_n^y}$. Under free entry, profits after corporate taxes equal the entry costs and therefore there are no dividends; in that case, capital owners only obtain income from land. The base for the sales tax in (A.17) is the total personal consumption expenditure of workers and capital owners defined in the previous section.

**Trade Imbalances** Three reasons give rise to differences between aggregate expenditures $P_n Q_n$ and sales $X_n$ of state $n$, and therefore create trade imbalances. First, differences in the ownership rates $\omega_n$ lead to differences between the gross domestic product of state $n$, $GDP_n$, and the gross income of residents of state $n$, $GSI_n$. Second, differences in ownership rates $\omega_n$ and in sales-apportioned corporate taxes $t_n^x$ across states create differences between the corporate tax revenue raised by state $n$'s government ($R_n^{corp}$) and the corporate taxes paid by residents of state $n$ ($TP_n^{corp}$). Third, there may be differences between taxes paid by residents of state $n$ to the federal government ($T_{n,fed}$) and the expenditures made by the federal government in state $n$ in either transfers to the state government in $n$ ($T_n^{fed\to st}$) or purchases of the final good produced in state $n$ ($G_{n,fed}$). As a result, the trade imbalance in state $n$, defined as difference between expenditures and sales in that state, can be written as follows:[1]

$$P_n Q_n - X_n = (GSI_n - GDP_n) + (R_n^{corp} - TP_n^{corp}) + \left(P_n G_{n,fed} + T_n^{fed\to st} - T_{n,fed}\right). \tag{A.18}$$

Letting $R = \sum_n r_n H_n$ and $\tilde{\Pi} = \sum_n \tilde{\Pi}_n$ be the pre-tax returns to the national portfolio of fixed factors and firms, we can rewrite some of the components of (A.18) as follows:[2]

$$GSI_n - GDP_n = \left(1-\chi^{FE}\right)\omega_n\left(\tilde{\Pi}-\tilde{\Pi}_n\right) + \omega_n R - r_n H_n, \tag{A.19}$$

$$R_n^{corp} = \frac{1}{\sigma}\left(t_n^x \frac{P_n Q_n}{X_n} + t_n^l\right) X_n, \tag{A.20}$$

$$TP_n^{corp} = b_n \sum_{n'} \left(\bar{t}_{n'} - t_{fed}^{corp}\right)\tilde{\Pi}_{n'}. \tag{A.21}$$

Replacing (A.19) to (A.21) into (A.18), and using (A.8) and (A.9) to express land payments and pre-tax profits as function of sales, after some manipulations we obtain:

$$\begin{aligned}\frac{P_n Q_n}{X_n} &= \frac{1}{\sigma - t_n^x}\left((\sigma-1)(1-\beta_n\gamma_n) + t_n^l + \frac{P_n G_{n,fed} + T_n^{fed\to st} - T_{n,fed}}{\tilde{\Pi}_n}\right) \\ &+ \frac{1}{\sigma - t_n^x}\left(\chi^{FE}\left(1-\beta_n\gamma_n\left(1-\bar{t}_n\right)\right) + \frac{\omega_n}{\tilde{\Pi}_n / \left(\Pi + R + \left(t_{fed}^{corp} - \chi^{FE}\right)\tilde{\Pi}\right)}\right)\end{aligned} \tag{A.22}$$

Expression (A.22) is used in the calibration to back out the ownership shares $\{\omega_n\}$ from observed data on trade imbalances. To implement it, we assume that transfers from the federal government to the state government in $n$ are entirely financed with federal taxes paid by residents of state $n$. Then, the ownership shares can be expressed as a function of other parameters and observables as follows:

$$\omega_n = \frac{\tilde{\Pi}_n}{\Pi + R + \left(t_{fed}^{corp} - \chi^{FE}\right)\tilde{\Pi}}\left[\left(\sigma - t_n^x\right)\left(\frac{P_n Q_n}{X_n}\right) - (\sigma-1)(1-\beta_n\gamma_n) - t_n^l - \chi^{FE}\left(1-\beta_n\gamma_n\left(1-\bar{t}_n\right)\right)\right]. \tag{A.23}$$

## B.3 General Equilibrium Conditions

We note that, using the definition of import shares in (A.11), imposing expression (17) for final good prices in every state is equivalent to imposing that expenditures shares in every state add up to 1.

$$\sum_n \lambda_{in} = 1 \text{ for all } i. \tag{A.24}$$

Additionally, by definition, aggregate sales by firms located in state $i$ are:

$$X_i = \sum_n \lambda_{ni} P_n Q_n. \tag{A.25}$$

[1]To reach this relationship, first impose goods market clearing (18) to obtain $P_n Q_n = P_n\left(C_n + G_{n,fed} + G_n + I_n\right)$. Then, note that personal-consumption expenditures can be written as $P_n C_n = GSI_n - (R_n^y + R_n^c + TP_n^{corp}) - T_{n,fed}$, where the terms between parentheses are tax payments made by residents of state $n$ to state governments and $T_{n,fed}$ are taxes paid to the federal government. Combining these two expressions and using the state's government budget constraint (28) gives $P_n Q_n = (GDP_n + P_n I_n) + (GSI_n - GDP_n) + (R_n^{corp} - TP_n^{corp}) + \left(P_n G_{n,fed} + T_n^{fed\to st} - T_{n,fed}\right)$. Adding and subtracting $GDP_n$ and noting that by definition $GDP_n = X_n - P_n I_n$ gives (A.18).

[2]Equations (A.19) and (A.21) hold by definition. For (A.20), combine (A.15) with (A.25) and (A.9).

This is equivalent to imposing that sales shares from every state add up to 1:

$$\sum_i s_{in} = 1 \text{ for all } n. \tag{A.26}$$

After several manipulations of the equilibrium conditions (available upon request), these shares can be expressed as functions of employment shares, wages, aggregate variables, and parameters as follows:

$$\lambda_{in} = A_{in} \left(\frac{w_n}{\bar{\pi}}\right)^{1-\kappa_1} \left(L_n h_n E_n\left[z\right]\right)^{1-\kappa_{2n}} \left(\frac{w_i}{\bar{\pi}}\right)^{\sigma-1} \left(L_i h_i E_i\left[z\right]\right)^{-\kappa_{3i}}, \tag{A.27}$$

$$s_{in} = \lambda_{in} \frac{P_i Q_i}{X_n}, \tag{A.28}$$

where $A_{in}$ is given by

$$A_{in} = \Theta_n \left( \frac{z_n^A}{\tau_{in}^A} \left(\frac{Z_i u_i^A}{v}\right)^{\frac{1}{1-\alpha_{W,i}}} \left(\frac{Z_n u_n^A}{v}\right)^{\frac{1-\gamma}{1-\alpha_{W,n}}\left(\frac{(\sigma-1)\alpha_F\chi_F-1}{\sigma-1}\chi^{FE}-1\right)} \right)^{\sigma-1}, \tag{A.29}$$

where $\left\{z_n^A, \tau_{in}^A, u_n^A\right\}$ are defined in (30) to (32) in the text, where $Z_n$ summarizes the impact of hours worked and skill heterogeneity,

$$Z_n = \frac{\zeta_n}{\zeta_n - (1-b_n)(1-\alpha_{W,n})} \left(\frac{\zeta_n}{\zeta_n - 1} z_{L,n}\right)^{1/\varepsilon_W + \alpha_{W,n}\chi_W} h_n^{(1-b_n)(1-\alpha_{W,n})} e^{-\alpha_h \frac{h_n^{1+1/\eta}}{1+1/\eta}},$$

and where $\Theta_n$ is a state-specific constant,

$$\begin{aligned} \Theta_n \equiv & \left(1+t_{fed}^w\right)^{1-(\sigma-1)\left(\frac{1-\chi^{FE}}{\varepsilon_F}+\alpha_F\chi_F\right)+\gamma\left(-(\sigma-1)+((\sigma-1)\alpha_F\chi_F-1)\chi^{FE}\right)} \\ & * \left(\frac{1-\beta}{\beta} H_n\right)^{\beta\gamma\left((\sigma-1)-[(\sigma-1)\alpha_F\chi_F-1]\chi^{FE}\right)} \left( \frac{f_{E,n}^{\chi^{FE}\left(\alpha_F - \frac{1}{\sigma-1}\right)\frac{\sigma-1}{\sigma}}}{\left((1-\beta)\gamma\right)^{\frac{1}{\sigma-1}-\frac{1-\chi^{FE}}{\varepsilon_F}-\alpha_F\chi_F}} \right)^{\sigma-1}. \end{aligned}$$

The parameters $\{\kappa_1, \kappa_{2n}, \kappa_3\}$ in (A.27) and (A.28) are given by:

$$\kappa_1 = (\sigma-1)\left(1+\alpha_F\chi_F+\frac{1-\chi^{FE}}{\varepsilon_F}\right) - \left((\sigma-1)\alpha_F\chi_F - 1\right)\chi^{FE}, \tag{A.30}$$

$$\begin{aligned} \kappa_{2n} = & (\sigma-1)\left(\frac{1-\chi^{FE}}{\varepsilon_F} + \alpha_F\chi_F + \gamma\beta - \frac{1+\varepsilon_W\chi_W\alpha_{W,n}}{\varepsilon_W(1-\alpha_{W,n})}(1-\gamma_n)\right) \\ & - \chi^{FE}\left(\gamma\beta\left((\sigma-1)\alpha_F\chi_F-1\right) - \frac{1+\varepsilon_W\chi_W\alpha_{W,n}}{\varepsilon_W}\frac{1-\gamma}{1-\alpha_{W,n}}\left((\sigma-1)\alpha_F\chi_F-1\right)\right), \end{aligned} \tag{A.31}$$

$$\kappa_{3i} = (\sigma-1)\frac{1+\varepsilon_W\chi_W\alpha_{W,i}}{\varepsilon_W(1-\alpha_{W,i})}. \tag{A.32}$$

As in the previous sections of this appendix, we let $\chi^{FE}$ be an indicator variable that equals 1 when we assume free entry of homogeneous firms and zero when we assume free mobility of heterogeneous firms.

Equations (A.24) to (A.29), together with (8) and (A.22), give the solution for import shares $\{\lambda_{in}\}$, export shares $\{s_{in}\}$, employment shares $\{L_n\}$, wages relative to average profits $\{w_n/\bar{\pi}\}$, government sizes $\{P_nG_n\}$, relative trade imbalances $\{P_nQ_n/X_n\}$, and utility $v$. The endogenous variables not included in this system can be recovered using the remaining equilibrium equations of the model.

## B.4 Uniqueness in a Special Case

Consider a special case of baseline the model with a fixed mass of ex-ante heterogeneous firms (i.e. $\chi^{FE}=0$) in which there is no dispersion in sales-apportioned corporate taxes across states ($t_n^x = t^x$ for all $n$), no cross-ownership of assets across states, and same preference for government spending across states ($\alpha_{W,n} = \alpha_W$). In this case, the adjusted amenities and productivities $u_n^A$ and $z_n^A$ defined in (32) and (30) become exogenous functions of fundamentals and own-state taxes. It is then possible to show that Conditions 1 to 3 and 4' of Allen et al. (2014) are satisfied (proof available upon request) and that, applying their Corollary 2, a sufficient uniqueness condition for the system of equations in $\{L_n, w_n/\bar{\pi}, v\}$ in (A.24) to (A.26) is

$$\frac{\sigma-(1-\kappa_3)}{\sigma(1-\kappa_2)-(1-\kappa_3)(1-\kappa_1)}>1, \tag{A.33}$$

$$\frac{\kappa_1-\kappa_2}{\sigma(1-\kappa_2)-(1-\kappa_3)(1-\kappa_1)}>1, \tag{A.34}$$

where $\kappa_1$ to $\kappa_3$ are defined in (A.30) to (A.32).

## B.5 General Equilibrium in Relative Changes

To perform counterfactuals, we solve for the changes in model outcomes as function of changes in taxes. Consider computing the effect of moving from the current distribution of state taxes, $\left\{t_n^y, t_n^c, t_n^x, t_n^l\right\}_{n=1}^N$ to a new distribution $\{(t_n^y)', (t_n^c)', (t_n^x)', (t_n^l)'\}_{n=1}^N$. As we discussed in Section 4.8, implementing counterfactuals in our framework requires simultaneously accounting for a mapping from changes in adjusted fundamentals to changes in outcomes and for a mapping from changes in taxes and in general-equilibrium outcomes to changes in adjusted fundamentals. The first mapping is given by (A.35) to (A.42) below, and the second is given by (A.43) to (A.45).

Defining $\hat{x}=x'/x$ as the counterfactual value of $x$ relative to its initial value, we have that the changes in import shares, export shares, number of workers, and wage per efficiency unit $\left\{\hat{\lambda}_{in}, \hat{s}_{in}, \hat{L}_n, \hat{w_n}\right\}$, as well as the welfare change of workers $\hat{v}$ must be such that conditions (A.24) and (A.26) hold:

$$\sum_n \lambda_{in}\hat{\lambda_{in}}=1 \text{ for all } i, \tag{A.35}$$

$$\sum_i s_{in}\hat{s_{in}}=1 \text{ for all } n, \tag{A.36}$$

where, using (A.27) and (A.28),

$$\hat{\lambda_{in}}=\hat{A_{in}}\left(\frac{\hat{w_n}}{\bar{\pi}}\right)^{1-\kappa_1}\left(\hat{h_n}\hat{L_n}\right)^{1-\kappa_{2n}}\left(\frac{\hat{w_i}}{\bar{\pi}}\right)^{\sigma-1}\left(\hat{h}_i\hat{L_i}\right)^{-\kappa_{3i}}, \tag{A.37}$$

$$\hat{s_{in}}=\hat{\lambda_{in}}\left(\frac{\widehat{P_iQ_i}}{X_i}\right)\frac{\hat{X}_i}{\hat{X}_n}, \tag{A.38}$$

where using (A.29),

$$\hat{A}_{in}=\left(\frac{\hat{z_n^A}}{\hat{\tau_{in}^A}}\left(\frac{\hat{Z}_i\hat{u_i^A}}{\hat{v}}\right)^{\frac{1}{1-\alpha_{W,i}}}\left(\frac{\hat{Z}_n\hat{u_n^A}}{\hat{v}}\right)^{\frac{1-\gamma_n}{1-\alpha_{W,n}}\left(\frac{(\sigma-1)\alpha_F\chi_F-1}{\sigma-1}\chi^{FE}-1\right)}\right)^{\sigma-1}, \tag{A.39}$$

where the impact of changes in hours worked and the skill distribution within each state is captured by

$$\hat{Z_n}=\left(\left(\hat{h_n}\right)^{1-\left(b_n^y\right)'}e^{-\frac{b_n^y-\left(b_n^y\right)'}{1+1/\eta}}\right)^{1-\alpha_{W,n}}\left(h_n^{b_n^y-\left(b_n^y\right)'}\right)^{1-\alpha_{W,n}}\frac{\zeta_n-(1-b_n^y)(1-\alpha_{W,n})}{\zeta_n-\left(1-(b_n^y)'\right)(1-\alpha_{W,n})} \tag{A.40}$$

and where, from (13), the change in the number of hours worked is

$$\hat{h_n}=\left(\frac{1-(b_n^y)'}{1-b_n^y}\right)^{\frac{1}{1+1/\eta}}. \tag{A.41}$$

Additionally, labor shares must add up to 1 :

$$\sum L_n\hat{L_n}=1. \tag{A.42}$$

From (30) to (32), the changes in the adjusted fundamentals are

$$\hat{\tau_{in}^A} = \frac{\sigma - \tilde{t}_{in}}{\sigma - \left(\tilde{t}_{in}\right)'}, \tag{A.43}$$

$$\hat{z_n^A} = \left(\frac{\left(1-(\bar{t}_n)'\right)/\left(\sigma-1+\chi^{FE}\left(1-(\bar{t}_n)'\right)\right)}{\left(1-\bar{t}_n\right)/\left(\sigma-1+\chi^{FE}\left(1-\bar{t}_n\right)\right)}\right)^{\frac{1}{\sigma-1}-\left(\frac{1-\chi^{FE}}{\varepsilon_F}+\alpha_F\chi_F\right)} \hat{G_n}^{\alpha_F}, \tag{A.44}$$

$$\hat{u_n^A} = \left(\frac{1-T_n'\left(w_n z_n^L \hat{w_n}\right)}{1-T_n\left(w_n z_n^L\right)}\frac{1+t_n^c}{1+(t_n^c)'}\right)^{1-\alpha_W}\left(\hat{G_n}\right)^{\alpha_{W,n}}. \tag{A.45}$$

where

$$\frac{1-T_n'\left(w_n z_n^L \hat{w_n}\right)}{1-T_n\left(w_n z_n^L\right)} = \hat{a_n^y}\frac{1+t_n^c}{1+(t_n^c)'}\hat{w_n}^{-\left(b_n^y\right)'}\left(w_n z_n^L\right)^{-\left(\left(b_n^y\right)'-b_n^y\right)}. \tag{A.46}$$

The variables $\left\{\frac{P_n\hat{Q}_n}{X_n}, \hat{G_n}, T_n', (\bar{t}_n)', (\tilde{t}_{in})'\right\}_{n=1}^N$ entering in (A.43) to (A.45) can be expressed as function of the original taxes $\left\{t_n^y, t_n^c, t_n^x, t_n^l\right\}_{n=1}^N$, the new tax distribution $\{(t_n^y)', (t_n^c)', (t_n^x)', \left(t_n^l\right)'\}_{n=1}^N$, and the new export shares $\{\hat{s_{in}} s_{in}\}_{n,i=1}^N$ using (9), (22), (24), (A.22), and (28). Hence, these equations, together with (A.35) to (A.42), give the solution for $\left\{\hat{\lambda}_{in}, \hat{s}_{in}, \hat{L}_n, \hat{w_n}\right\}$ and $\hat{v}$. The new government sizes and trade deficits also depend on the new values of $\tilde{\Pi}$ and $\Pi + R$; these variables can be expressed as a function of initial conditions and changes in the endogenous variables.

# C Appendix to Section 5

**Proof of Proposition 1** Under the assumptions in the proposition, the efficient allocations follow from optimization of the following Lagrangian:

$$\begin{aligned}\mathcal{L} = v &- \sum_n \lambda_{1n}\left[v - U_n\left(\frac{C_n^L}{L_n}, G_n\right)\right] - \sum_n \lambda_{2n}\left[v_n^K - U_n^K\left(\frac{C_n^K}{K_n}, G_n\right)\right] \\ &- \lambda_3\left(\sum_n C_n^L + \sum_n C_n^K + \sum_n G_n + \sum_n I_n - \sum_n F_n\left(L_n, I_n\right)\right) \\ &- \lambda_4\left(\sum L_n - 1\right).\end{aligned} \tag{A.47}$$

The efficient allocations result from maximizing the welfare of workers $v$ given arbitrary levels of welfare of capital owners, $v_n^K$. The first term in square brackets in the first line is the spatial mobility constraint, where $U_n(c,g)$ is the direct utility function defined in (2) under the assumption of no disutility from labor, and where $U_n^K(c,g)$ is the utility of each capital owner in $n$. The second line shows the goods feasibility constraint, where

$$F_n\left(L_n, I_n\right) = z_n^0\left[\frac{1}{\gamma_n}\left(\frac{H_n}{\beta_n}\right)^{\beta_n}\left(\frac{L_n}{1-\beta_n}\right)^{1-\beta_n}\right]^{\gamma_n}\left(\frac{I_n}{1-\gamma_n}\right)^{1-\gamma_n}$$

is the production technology. The last line of (A.47) is national labor market clearing. Except for the existence of intermediates, the arbitrary many regions, and the immobile capital owners, (A.47) is the same optimization problem considered in Flatters et al. (1974) and Wildasin (1980). Letting $U_{nc} \equiv \partial U_n(c,g)/\partial c$ and $F_{nL} \equiv \partial F_n/\partial L_n$, taking the first order condition over $L_n$ and $C_n$ we obtain

$$[L_n] \qquad \lambda_3 F_{nL} = \lambda_4 + \frac{C_n^L}{L_n}\lambda_{1n}\frac{U_{nc}^L}{L_n},$$

$$\left[C_n^L\right] \qquad \lambda_{1n}\frac{U_{nc}^L}{L_n} = \lambda_3.$$

Combining these two conditions we obtain $F_{nL} - C_n^L/L_n = \lambda_4/\lambda_3$. Under $t_{fed}^w = 0$ the market allocation gives $w_n = F_{nL}$. Absent compensating differentials, mobility of workers implies that $C_n^L/L_n$ is constant across locations, which gives part i) of the proposition. More generally, we have that $w_n - C_n^L/L_n$, which equals tax payments in the decentralized equilibrium, is equalized across locations, which gives part ii).

**Proof of Proposition 2** Consider a tax structure with only state sales and income taxes. Assume no trade costs ($\tau_{in} = 1$ for all $i, n$), perfect substitutability across varieties ($\sigma \to \infty$), homogeneous firms ($\varepsilon_F \to \infty$), and constant labor supply ($\zeta_n \to \infty$ and $h_n$ constant). Because goods are perfect substitutes ($\sigma \to \infty$) and there are no trade costs ($\tau_{in} = 1$) the production cost $c_n$ must be equalized across regions, and normalized to 1. This must also be the price of the final good produced everywhere. Because firms are homogeneous ($\varepsilon_F \to \infty$), it follows from (27) that the summary statistic of the productivity distribution in $n$ equals the common component of productivity, $\tilde{z}_n = z_n^0$. Using (A.6), total production in region $n$ is

$$\left(\frac{z_n^0}{\gamma_n}\right)^{1/\gamma_n} \left(\frac{H_n}{\beta_n}\right)^{\beta_n} \left(\frac{L_n}{1-\beta_n}\right)^{1-\beta_n}. \tag{A.48}$$

From (4), state-specific appeal is:

$$v_n = u_n \left(\frac{G_n}{L_n^{\chi_W}}\right)^{\alpha_{W,n}} \left(\left(1-T_n\right) w_n\right)^{1-\alpha_{W,n}}. \tag{A.49}$$

From (A.7), labor demand in state $n$ is given by the condition that labor costs equal the marginal product of labor, $w_n = MPL_n$, given by

$$MPL_n = Z_{n,0} L_n^{-\beta_n}, \tag{A.50}$$

where $Z_{n,0} = \left(1-\beta_n\right)^{\beta_n} \beta_n^{-\beta_n} \left(z_n^0/\gamma_n\right)^{1/\gamma_n} H_n^{\beta_n}$. Labor supply in $n$ follows from (7). Equating local labor demand and local labor supply gives the solution for employment in $n$,

$$L_n^*\left(v\right) = \left(\frac{\left(Z_n\left(1-T_n\right)\right)^{1-\alpha_{W n}}}{v}\right)^{\frac{1}{1/\varepsilon_W + \alpha_{W,n}\chi_W + \left(1-\alpha_{W,n}\right)\beta_n}} \tag{A.51}$$

where $Z_n = Z_{n,0}\left(u_n G_n^{\alpha_W}\right)^{\frac{1}{1-\alpha_W}}$. National labor-market clearing then gives the solution for worker welfare $v$ as the value where $H^*\left(v\right) \equiv \sum_{n=1}^N L_n^*\left(v\right) = 1$. $H^*\left(v\right)$ is decreasing in $v$ so that there can only be a unique solution for $v$. Assume now that $\alpha_{W,n} = \alpha_W$ for all $n$. Then, letting

$$\zeta = \frac{1-\alpha_W}{1/\varepsilon_W + \alpha_W\chi_W + \left(1-\alpha_W\right)\beta} > 0,$$

the solution for worker welfare is:

$$v = \left(\sum_n \left(Z_n\left(1-T_n\right)\right)^{\zeta}\right)^{\frac{1-\alpha_W}{\zeta}}. \tag{A.52}$$

Let $v'$ be welfare under a distribution of taxes where every tax rate is brought to the mean of the initial distribution, $T_n' = N^{-1}\sum T_n$ for all $n$. Then, $v' > v$ if and only if

$$E[Z_n^{\zeta}](E[1-T_n])^{\zeta} > cov[Z_n^{\zeta}, (1-T_n)^{\zeta}] + E[Z_n^{\zeta}]E[(1-T_n)]^{\zeta} \tag{A.53}$$

where $E$ and $cov$ denote the sample mean and covariance. This expression can be rearranged to reach

$$\frac{\mathbb{E}\left[1-T_n\right]^{\zeta} - \mathbb{E}\left[\left(1-T_n\right)^{\zeta}\right]}{sd\left(\left(1-T_n\right)^{\zeta}\right)} > cv\left(Z_n^{\zeta}\right) corr\left[Z_n^{\zeta}, \left(1-T_n\right)^{\zeta}\right] \tag{A.54}$$

where $cv$ and $sd$ denote the coefficient of variation and the standard deviation. Therefore, $v' > v$ if $corr[Z_n^{\zeta}, (1-T_n)]^{\zeta}$ is low enough, and $v' > v$ if $corr[Z_n^{\zeta}, (1-T_n)]^{\zeta}$ is large enough. Part i) follows from the fact that $\zeta = 1/\beta$ when $\varepsilon_W \to \infty$ and $\chi_W = 0$ . Part ii) follows from the example in the body of the text.

# D Appendix to Section 6

## D.1 Model-implied Fundamentals

This section shows the composite term $A_{in}$, is related to measures of local amenities and market access. We recover measures of the composite terms $A_{in}$ from observed data and estimated parameters using equation (A.27). Then, we relate the composite terms $A_{in}$ to its determinants in the model. We focus on testing the predictions of equation (A.29) for the relationship between $A_{in}$ and observable exogenous determinants of trade costs $\tau_{in}^A$ and amenities in states $i$ and $n$, $u_i^A$ and $u_n^A$ by estimating regressions of the form

$$\ln A_{in} = b_0 + b_1 \ln \tau_{in}^A + b_2 u_i^A + b_3 u_n^A + e_{in},$$

where we use distance as our proxy for trade costs $\tau_{in}^A$ and data on measures of temperature and air quality in a state as our proxy for $u_i^A$ and $u_n^A$. Table A.5 reports the results of these regressions. We find three main results: First, there is a negative relation between distance between states and the composite term $A_{in}$. Second, there is a positive relation between $A_{in}$ and observable covariates that increase state $i$'s exogenous amenity level $u_i^A$ (and vice versa). Column 2 shows that states with a higher minimum temperature, and with lower maximum temperatures and precipitation have higher values of $A_{in}$. We also find that states with a lower number of toxic sites have higher values of $A_{in}$, but this relation, as well as the relations with measures of air quality, are not statistically significant. Finally, there is a negative relation between $A_{in}$ and observable covariates that increase state $n$ exogenous amenity level $u_n^A$ (and vice versa). Column 3 shows that destination states with more amenable weather (higher minimum temperature, lower maximum temperatures, and less precipitation) have lower values of $A_{in}$. These relationships are consistent with (A.29). The only relation that contradicts the prediction of the model is that of particulate matter in destination states, which may reflect other factors like the level of economic activity. Overall, these results provide evidence that the model-implied fundamentals have sensible empirical foundations.

Table A.5: Regressions of $\ln(A_{in})$ on Own-State and Other-State Amenities

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log Distance in Miles | -1.467*** | -1.296*** | -1.506*** | -1.298*** |
| | (0.292) | (0.170) | (0.313) | (0.174) |
| Min Temp (Origin) | | 1.422*** | | 1.418*** |
| | | (0.481) | | (0.483) |
| Max Temp (Origin) | | -0.988** | | -0.985** |
| | | (0.379) | | (0.380) |
| Precipitation (Origin) | | -0.003* | | -0.003* |
| | | (0.002) | | (0.002) |
| Toxic Site (Origin) | | -0.269 | | -0.268 |
| | | (0.538) | | (0.539) |
| Particulate Matter (Origin) | | 0.354 | | 0.352 |
| | | (0.285) | | (0.286) |
| Ozone Days (Origin) | | 0.069 | | 0.070 |
| | | (0.059) | | (0.060) |
| Min Temp (Destination) | | | -0.203*** | -0.096*** |
| | | | (0.014) | (0.031) |
| Max Temp (Destination) | | | 0.070*** | -0.010 |
| | | | (0.006) | (0.018) |
| Precipitation (Destination) | | | 0.001*** | 0.001*** |
| | | | (0.000) | (0.000) |
| Toxic Site (Destination) | | | 0.075*** | 0.044** |
| | | | (0.015) | (0.021) |
| Particulate Matter (Destination) | | | -0.103*** | -0.061** |
| | | | (0.028) | (0.024) |
| Ozone Days (Destination) | | | 0.035*** | 0.034*** |
| | | | (0.005) | (0.007) |
| Observations | 2122 | 2099 | 2114 | 2091 |

NOTES: The table reports results of a regression of the form: $\ln A_{in} = b_0 + b_1 \ln distance_{in} + b_2 u_i^A + b_3 u_n^A + e_{in}$, where $A_{in}$ is constructed from the data using equation (A.27), and where $u_i^A$ and $u_n^A$ are measures of amenities in origin and destination states. We estimate these regressions using a cross-section of data, where data on amenities comes from Couture et al. (2018). Amenity data are population-weighted averages at the state level. All models allow for clustered standard errors by the origin state and the destination state. Standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

## D.2 Appendix to Section 6.2

### D.2.1 Construction of Covariates in Worker and Firm Mobility Equation

We first describe how we construct the variable $z_n^L w_{nt}$ entering the covariate $\tilde{y}_{nt}$ in (34) and, through (A.46), the system of equilibrium equations in changes used for counterfactuals. In the model, the hourly wage of a worker $l$ in state $n$ is $w_n^h(l) = z_n^l w_n$, where $w_n$ is the wage per efficiency unit. Given the assumption that distribution of efficiency units within each state is Pareto with parameters $\left(z_n^L, \zeta_n\right)$, average hourly income per worker in state $n$ is $\mathbb{E}_n\left[w_n^h(l)\right] = z_n^L \frac{\zeta_n}{\zeta_n - 1} w_n$. Assuming that the shape of the Pareto distribution $\zeta_n$ is constant over time, we obtain

$$z_n^L w_{nt} = \mathbb{E}_n\left[w_{nt}^h(l)\right] \frac{\zeta_n - 1}{\zeta_n}, \tag{A.55}$$

where $\mathbb{E}_n\left[w_{nt}^h(l)\right]$ is empirically measured as the average hourly wage across individuals living in state $n$ in year $t$. Using again the assumption that the distribution of efficiency units across workers within a state is Pareto, the shape parameter $\zeta_n$ can be estimated using information on the average and variance of the distribution of hourly wages across workers living in state $n$:

$$(\zeta_n - 2)\zeta_n = \frac{\mathbb{E}_n\left[w_{nt}\left(l\right)\right]^2}{\mathbb{V}_n\left[w_{nt}\left(l\right)\right]}. \tag{A.56}$$

For each state $n$ and period $t$, we construct $A_{nt}^S$ using the information on the estimated $\zeta_n$ and estimated progressivity parameter $b_{nt}^y$. See Appendix F.1 for detailed information about the construction of income tax schedule parameters.

To construct measures of after-tax real earnings $\tilde{y}_{nt}$, market potential $MP_{nt}$, real government services $\tilde{R}_{nt}$, and unit costs $c_{nt}$, we need data on prices. We use the consumer price index from the Bureau of Labor Statistics. This is the same price data that is used in the estimation of the labor equation to construct measures of real government spending and real wages.

Constructing unit costs also requires data on the price of structures $r_{nt}$, which is is not available at an annual frequency. To avoid this shortcoming in the data, we construct an annual series of unit costs by setting the local price of structures to equal the local price index, resulting in the following measure of unit costs: $c_{nt} = (w_{nt}^{1-\beta_n} P_{nt}^{\beta_n})^{\gamma_n} P_{nt}^{1-\gamma_n}$.[3]

To construct measures of $\{\bar{t}_{nt}\}_{n=1}^N$ and $\{\tilde{t}_{n'nt}\}_{n=1,n'=1}^{N,N}$ (which enters the market potential $MP_{nt}$), we need information on the share of total sales generated in state $n$ that accrue to state $n'$. Annual data on trade flows across U.S. states does not exist. To overcome this data limitation, we set export shares in any period $t$ equal to the average of the recorded export shares for the years 1993 and 1997, i.e., $s_{int} = 0.5 \times (s_{in,1993} + s_{in,1997})$ . We also use the same information on export shares to construct a proxy for the term $\{\tau_{n'nt}\}_{n=1,n'=1}^{N,N}$ entering the expression for $\{MP_{nt}\}_{n=1}^N$. Specifically, we set $\tau_{n'nt} = dist_{n'n}^{\zeta}$, where $\zeta = 0.8/(\sigma - 1)$ and 0.8 is the point estimate of the elasticity of cross-state export shares with respect to distance, controlling for year, exporter and importer fixed effects.

We also need information on total state expenditures $\{P_{nt}Q_{nt}\}_{n=1}^N$ to a measure for $\{MP_{nt}\}_{n=1}^N$. Since expenditures are not observed in every year, we follow the predictions of the model and construct a proxy for $P_{nt}Q_{nt}$ for every state $n$as a function of state $n$'s GDP by combining (A.7), (A.12), and (A.22) to obtain

$$P_{nt}Q_{nt} = \frac{(\sigma - 1)(1 - \beta_n\gamma_n) + a_{nt} + t_n^l}{\sigma - t_n^x} \frac{\sigma}{\gamma_n(\sigma - 1) + 1} GDP_{nt}, \tag{A.57}$$

where $a_{nt} \equiv b_n(\Pi + R + t_{fed}^{corp}\tilde{\Pi})(\tilde{\Pi}_n)^{-1}$. State GDP is observed in every year, but $a_{nt}$ is not. Hence, to compute a yearly measure of $P_{nt}Q_{nt}$, we set its value to that observed in the calibration: $a_{nt} = a_{n,2007}$ for all $t$.[4]

## D.3 Construction of Instrument for Market Potential

We define the instrument $MP_{nt}^*$ as a variable that has a similar structure to market potential $MP_{nt}$ in (41) but that differs from it in that we substitute the components $E_{nt}$, $P_{nt}$, and $\tilde{t}_{n'nt}$ that might potentially be correlated with $\nu_{nt}^M$ with functions of exogenous covariates that we respectively denote as $E_{nt}^*$, $P_{nt}^*$, and $\tilde{t}_{n'nt}^*$ :

$$MP_{nt}^* = \sum_{n' \neq n} E_{n't}^* \left(\frac{\tau_{n'nt}}{P_{n't}^*} \frac{\sigma}{\sigma - \tilde{t}_{n'nt}^*} \frac{\sigma}{\sigma - 1}\right)^{1-\sigma}. \tag{A.58}$$

To implement this expression, we need to construct measures of the variables $E_{nt}^*$, $P_{nt}^*$, and $\tilde{t}_{n'nt}^*$. We construct $E_{nt}^*$ using (A.57) with lagged GDP instead of period $t's$ GDP:

$$E_{nt}^* = \frac{(\sigma - 1)(1 - \beta_n\gamma_n) + a_{nt} + t_n^l}{\sigma - t_n^x} \frac{\sigma}{\gamma_n(\sigma - 1) + 1} GDP_{n,t-1}$$

We set $P_{n,t}^* = 1 + t_{n,t}^c$. We construct $\tilde{t}_{n'nt}^*$ using the expression for $\tilde{t}_{ni}$ in (24) evaluated at hypothetical export shares defined as relative inverse log distances:

[3]Projecting the decadal data on rental prices $r_{nt}$ on wages and local price indices, $w_{nt}$ and $P_{nt}$, and using the projection estimates in combination with annual data on $w_{nt}$ and $P_{nt}$ to compute predicted rental prices, $\hat{r}_{nt}$, and predicted unit costs, $c_{nt} = (w_{nt}^{1-\beta_n}\hat{r}_{nt}^{\beta_n})^{\gamma_n} P_{nt}^{1-\gamma_n}$, produces similar estimates of the structural parameters $\varepsilon_F$ and $\alpha_F$.

[4]Using an alternate definition of $P_{nt}Q_{nt}$, i.e., $P_{nt}Q_{nt} = constant{*}GDP_{nt}$ where the constant is an OLS estimate of the derivative of total expenditures with respect to GDP in those years in which we observe both components, yields very similar results.

$$s^*_{int} = \frac{\ln(dist_{in})^{-1}}{\sum_{i\neq n}\ln(dist_{in})^{-1}+1} \quad \forall t, i\neq n \qquad \text{and} \qquad s^*_{iit} = \frac{1}{\sum_{i\neq n}\ln(dist_{in})^{-1}+1} \quad \forall t.$$

## D.4 Robustness Checks: Labor Supply

This section presents a series of robustness checks that address a number of potential concerns about our instrument choice and labor supply specification. Table A.6 presents GMM estimates for structural parameters when government spending is measured using actual, as opposed to model-based, tax revenue. Table A.7 reports estimates from a specification in which we use a wage Bartik instrument instead of a payroll Bartik instrument. In Table A.8 we estimate structural parameters in the case of no unobserved worker heterogeneity. In Table A.9 we ignore the intensive margin of labor supply.

First, Table A.6 reports GMM estimates for structural parameters of the labor supply equation when government spending $\tilde{R}_{nt}$ is measured using actual, as opposed to model-based, tax revenue.

Second, Table A.7 reports GMM estimates that differ from the baseline ones in Section 6.2 in that $\boldsymbol{Z}^B_{nt}$ contains a wage Bartik instrument instead of a payroll Bartik instrument:

$$\text{BtkW}_{nt} = \sum_k \frac{L_{kn,1974}}{L_{n,1974}} \frac{w_{kt}-w_{k,t-10}}{w_{k,t-10}},$$

where $w$ denotes real hourly wages.

Third, we consider a specification in which we do not account for unobserved worker heterogeneity. Specifically, Table A.8 shows GMM estimates from the following model:

$$\ln L_{nt} = a_{0,n}\ln\tilde{y}_{nt} + b_{0,n}\ln\tilde{R}_{nt} + \psi^L_t + \xi^L_n + \nu^L_{nt},$$

where the shape parameter of the distribution of efficiency units, $\zeta_{n,}$ is set to 1 and, as a consequence, the hourly wage adjusted for efficiency units is equal to the raw wage observed in the Current Population Survey (CPS).

Finally, Table A.9 reports worker parameter estimates from a specification in which we do not account for the intensive margin of labor supply. This implies that real after-tax earnings are defined as:

$$\tilde{y}_{nt} \equiv \frac{a^y_{nt}}{1+t^c_{nt}}\frac{1}{P_{nt}}\left(h_{nt}w^z_{nt}\right)^{1-b^y_{nt}}.$$

Table A.6: GMM Estimates of Worker Parameters: Tax Revenue Robustness

| Instruments | Restrictions on $\alpha_{W,n}$ | $\varepsilon_W$ | | $\alpha_W$ | |
|---|---|---|---|---|---|
| | | $\chi_W = 0$ | $\chi_W = 1$ | $\chi_W = 0$ | $\chi_W = 1$ |
| $\boldsymbol{Z}_{nt}^T$ | $\alpha_{W,n} = \alpha_W$ | 1.23*** | 1.96** | .3** | .3** |
| | | (.33) | (.96) | (.12) | (.12) |
| $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | 1.81*** | 3.54** | .27** | .27** |
| | | (.48) | (1.78) | (.12) | (.12) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | 1.26*** | 1.81*** | .24** | .24** |
| | | (.28) | (.69) | (.11) | (.11) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \frac{R_n}{GDP_n}$ | .72*** | 1.4*** | | |
| | | (.23) | (.33) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0.04 = \text{Mean}\frac{R_n}{GDP_n}$ | 1.1*** | 1.15*** | | |
| | | (.31) | (.34) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0$ | 1.03*** | 1.03*** | | |
| | | (.3) | (.3) | | |

Notes: This table shows the GMM estimates for structural parameters entering the labor mobility equation (33). The data are at the state-year level. Each column has 712 observations. Every specification includes state and year fixed effects. Observations are weighted using state population. The instrument vectors used to compute the estimates in each row are indicated under the heading "Instruments". Similarly, restrictions on $\alpha_{W,n}$ are described under the heading "Restrictions on $\alpha_{W,n}$". Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

Table A.7: GMM Estimates of Worker Parameters: Labor Market Bartik IV Robustness

| Instruments | Restrictions on $\alpha_{W,n}$ | $\varepsilon_W$ | | $\alpha_W$ | |
|---|---|---|---|---|---|
| | | $\chi_W = 0$ | $\chi_W = 1$ | $\chi_W = 0$ | $\chi_W = 1$ |
| $\boldsymbol{Z}_{nt}^T$ | $\alpha_{W,n} = \alpha_W$ | 1.42*** | 2.1*** | .23*** | .23*** |
| | | (.36) | (.8) | (.07) | (.07) |
| $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | .86* | 1.05 | .21* | .21* |
| | | (.47) | (.65) | (.13) | (.13) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | 1.17*** | 1.47*** | .18*** | .18*** |
| | | (.31) | (.5) | (.07) | (.07) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \frac{R_n}{GDP_n}$ | .47* | 1.3*** | | |
| | | (.25) | (.33) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0.04 = \text{Mean}\frac{R_n}{GDP_n}$ | .99*** | 1.03*** | | |
| | | (.3) | (.32) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0$ | .83*** | .83*** | | |
| | | (.27) | (.27) | | |

Notes: This table shows the GMM estimates for structural parameters entering the labor mobility equation (33). The data are at the state-year level. Each column has 712 observations. Every specification includes state and year fixed effects. Observations are weighted using state population. The instrument vectors used to compute the estimates in each row are indicated under the heading "Instruments". Similarly, restrictions on $\alpha_{W,n}$ are described under the heading "Restrictions on $\alpha_{W,n}$". Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: GMM Estimates of Worker Parameters: No Unobserved Worker Heterogeneity

| Instruments | Restrictions on $\alpha_{W,n}$ | $\varepsilon_W$ | | $\alpha_W$ | |
|---|---|---|---|---|---|
| | | $\chi_W = 0$ | $\chi_W = 1$ | $\chi_W = 0$ | $\chi_W = 1$ |
| $\boldsymbol{Z}_{nt}^T$ | $\alpha_{W,n} = \alpha_W$ | 1.42*** | 2.09*** | .23*** | .23*** |
| | | (.36) | (.79) | (.07) | (.07) |
| $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | 1.79*** | 2.25** | .11* | .11* |
| | | (.63) | (.93) | (.06) | (.06) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | 1.35*** | 1.72*** | .16*** | .16*** |
| | | (.3) | (.52) | (.06) | (.06) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \frac{R_n}{GDP_n}$ | .74*** | 1.48*** | | |
| | | (.23) | (.33) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0.04 = \text{Mean}\frac{R_n}{GDP_n}$ | 1.19*** | 1.25*** | | |
| | | (.32) | (.35) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0$ | 1.04*** | 1.04*** | | |
| | | (.3) | (.3) | | |

NOTES: This table shows the GMM estimates for structural parameters entering the labor mobility equation (33) when $\ln A_n^S = 0$ and $w_n^z = w_n^{CPS}$, i.e., there is no unobserved worker heterogeneity. The data are at the state-year level. Each column has 712 observations. Every specification includes state and year fixed effects. Observations are weighted using state population. The instrument vectors used to compute the estimates in each row are indicated under the heading "Instruments". Similarly, restrictions on $\alpha_{W,n}$ are described under the heading "Restrictions on $\alpha_{W,n}$". Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

Table A.9: GMM Estimates of Worker Parameters: No Intensive Margin of Labor Supply

| Instruments | Restrictions on $\alpha_{W,n}$ | $\varepsilon_W$ | | $\alpha_W$ | |
|---|---|---|---|---|---|
| | | $\chi_W = 0$ | $\chi_W = 1$ | $\chi_W = 0$ | $\chi_W = 1$ |
| $\boldsymbol{Z}_{nt}^T$ | $\alpha_{W,n} = \alpha_W$ | 1.48*** | 2.24*** | .23*** | .23*** |
| | | (.37) | (.86) | (.06) | (.06) |
| $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | 1.81*** | 2.3** | .12* | .12* |
| | | (.64) | (.95) | (.06) | (.06) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \alpha_W$ | 1.39*** | 1.8*** | .16*** | .16*** |
| | | (.31) | (.54) | (.06) | (.06) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_{W,n} = \frac{R_n}{GDP_n}$ | .8*** | 1.39*** | | |
| | | (.25) | (.29) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0.04 = \text{Mean}\frac{R_n}{GDP_n}$ | 1.2*** | 1.26*** | | |
| | | (.32) | (.36) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_W = 0$ | 1.03*** | 1.03*** | | |
| | | (.3) | (.3) | | |

NOTES: This table shows the GMM estimates for structural parameters entering the labor mobility equation (33) when $\eta \to 0$, i.e., the labor supply has no intensive margin responses. The data are at the state-year level. Each column has 712 observations. Every specification includes state and year fixed effects. Observations are weighted using state population. The instrument vectors used to compute the estimates in each row are indicated under the heading "Instruments". Similarly, restrictions on $\alpha_{W,n}$ are described under the heading "Restrictions on $\alpha_{W,n}$". Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

## D.5 Robustness Checks: Firm Mobility

Table A.10 presents a robustness check for the firm mobility equation. Using the baseline model in Section 6.2, we measure government spending $\tilde{R}_{nt}$ using model-based, as opposed to actual, tax revenue.

Table A.10: GMM Estimates of Firm Parameters: Tax Revenue Robustness

| Instruments | Restrictions on $\alpha_F$ | $\varepsilon_F$ | | $\alpha_F$ | |
|---|---|---|---|---|---|
| | | $\chi_F = 0$ | $\chi_F = 1$ | $\chi_F = 0$ | $\chi_F = 1$ |
| $\boldsymbol{Z}_{nt}^T$ | None | 2.5*** | 2.15*** | -.06* | -.06* |
| | | (.28) | (.27) | (.04) | (.04) |
| $\boldsymbol{Z}_{nt}^B$ | None | 2.74*** | 2.66*** | -.01 | -.01 |
| | | (.32) | (.33) | (.03) | (.03) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | None | 2.46*** | 2.3*** | -.03 | -.03 |
| | | (.26) | (.27) | (.03) | (.03) |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_F = 0.04 = \text{Mean}\frac{R_n}{GDP_n}$ | 2.29*** | 2.53*** | | |
| | | (.25) | (.31) | | |
| $\boldsymbol{Z}_{nt}^T$, $\boldsymbol{Z}_{nt}^B$ | $\alpha_F = 0$ | 2.45*** | 2.45*** | | |
| | | (.26) | (.26) | | |

NOTES: This table shows GMM estimates for structural parameters entering the firm mobility equation (40). Data are at the state-year level. Each column has 587 observations, which is lower than the worker estimation due to data requirements for constructing a measure of the market potential and unit costs terms (see Appendix D.3 for details). Every specification includes state and year fixed effects. The instrument vectors used to compute the estimates in each row are indicated under the heading "Instruments". Similarly, restrictions on $\alpha_F$ are described under the heading "Restrictions on $\alpha_F$". Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

## D.6 Supplemental: 2SLS Estimates of Worker Parameters

This section presents both Ordinary Least Squares (OLS) and Two-Stage Least Squares (2SLS) estimates for the auxiliary parameters $a_0$ and $a_1$ in (33). To implement a 2SLS estimator, we consider the simplified case in which there is no unobserved worker heterogeneity (i.e., the case in which $\ln A_n^S = 0$ and thus $w_n^z = w_n^{CPS}$). Appendix Table A.8 shows that the estimates in this case are nearly identical to the baseline estimates. When computing this 2SLS estimator, we use the two instrument vectors described in Section 6.2, $\boldsymbol{Z}_{nt}^T$ and $\boldsymbol{Z}_{nt}^B$, first separately and then jointly.

Table A.11 provides the estimates of the first-stage regression corresponding to the 2SLS estimation of $a_0$ and $a_1$. Column (1) shows the estimates of a regression of after-tax real wages on the instrument vector $\boldsymbol{Z}_{nt}^T$ and state and year fixed effects. Column (4) does the same for real government services $\tilde{R}_{nt}$. The coefficients on external taxes indicate that being "close" to high sales tax (and high sales-apportioned corporate tax) states tends to be associated with lower after-tax real wages. Real government services tend to be lower when the state is "close" to high income tax states. Columns (2) and (5) show the results using the Bartik instruments $\boldsymbol{Z}_{nt}^B$. Initial state-industry specific shares weighted national industry-specific payroll changes and initial state-type of tax specific shares weighted national tax revenue shocks tend to be associated with higher state earnings and government service provision. The first stage for earnings is a bit underpowered in the Bartik IV specification, whereas the state tax revenue first stage is fairly strong. Columns (3) and (6) show the first stage results when both sets of instruments are included. The F-statistics of joint significance of the instruments conditional on state and year fixed effects are 8.6 in column (3) and 13.4 in column (6). Additionally, the Cragg-Donald Wald F-statistic is 9.9 and the Kleibergen-Paap Wald F-statistic is 7.8.

As mentioned in the main text, our model predicts that OLS estimates of $a_0$ and $a_1$ are asymptotically biased due to the dependence of after-tax real earnings and government spending on unobserved amenities or government efficiency accounted for in the term $\nu_{nt}^L$. Specifically, our model predicts that amenities in a state are negatively correlated with its after-tax real earnings and positively correlated with its real government spending. Intuitively, higher amenities in a state attract workers, shift out the labor supply curve, and lower wages. This increase in the number of workers also raises the tax revenue and thus increases government spending. Our model thus predicts that the OLS estimate of $a_0$ is biased downwards, and the OLS estimate of $a_1$ is biased upwards. Therefore, if our

Table A.11: First Stage of Labor-Supply Equation

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | $\ln \tilde{y}_{nt}$ | | | $\ln \tilde{R}_{nt}$ | |
| | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ |
| $t^{*x}_{nt}$ | -0.39 | | -0.40 | 0.80 | | 0.36 |
| | (0.25) | | (0.25) | (0.74) | | (0.70) |
| $t^{*c}_{nt}$ | 3.28*** | | 3.19*** | 0.13 | | 0.37 |
| | (0.55) | | (0.55) | (1.93) | | (1.77) |
| $t^{*y}_{nt}$ | 0.48 | | 0.53 | -8.11*** | | -6.40*** |
| | (0.45) | | (0.46) | (1.40) | | (1.37) |
| $BtkP_{nt}$ | | 0.06** | 0.06** | | 0.01 | -0.01 |
| | | (0.03) | (0.02) | | (0.07) | (0.07) |
| $BtkTR_{nt}$ | | -0.03 | -0.01 | | 1.05*** | 0.90*** |
| | | (0.04) | (0.04) | | (0.19) | (0.20) |
| R-squared | 0.946 | 0.944 | 0.947 | 0.992 | 0.992 | 0.993 |
| F-stat | 12.1 | 3.4 | 8.6 | 12.6 | 15.3 | 13.4 |

NOTES: This table shows first-stage estimates for the labor mobility equation (33) when $\ln A^S_n = 0$ and $w^z_n = w^{CPS}_n$, i.e., there is no unobserved worker heterogeneity. The dependent variables are after-tax real earnings and real government expenditures in columns 1-3 and 4-6, respectively. Data are at the state-year level. Every specification includes state and year fixed effects. Each column has 712 observations. F-statistics refer to specifications that do not control for state and year dummies. Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

instrument vectors were to be valid, we should should obtain 2SLS estimates of $a_0$ and $a_1$ that are, respectively, higher and lower than their OLS counterparts.

Table A.12 presents OLS and 2SLS estimates of $a_0$ and $a_1$. Column (1) shows the OLS estimates. Columns (2)-(4) show the 2SLS estimates. Compared to the 2SLS estimates, the OLS estimates imply a lower elasticity of labor supply with respect to after-tax real earnings and a larger one with respect to real government spending. This difference between the OLS and the 2SLS estimates is consistent with our model's predictions. In addition, the 2SLS estimates that rely on different instrument vectors are quite similar. The implications of these reduced-form estimates of $a_0$ and $a_1$ for our structural parameters are shown at the bottom of the table.

Table A.12: OLS and 2SLS Estimates of Local Labor Supply Parameters

| | (1) OLS | (2) 2SLS | (3) 2SLS | (4) 2SLS |
|---|---|---|---|---|
| | | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ |
| $\ln \tilde{y}_{nt}$ | 0.28*** | 1.36*** | 1.58*** | 1.35*** |
| | (0.06) | (0.34) | (0.61) | (0.30) |
| $\ln \tilde{R}_{nt}$ | 0.44*** | 0.32*** | 0.20* | 0.23** |
| | (0.03) | (0.12) | (0.11) | (0.10) |
| Structural Parameters | | | | |
| $\varepsilon_W$ for $\chi_W = 0$ | .72*** | 1.67*** | 1.79*** | 1.59*** |
| | (.07) | (.39) | (.63) | (.34) |
| $\varepsilon_W$ for $\chi_W = 1$ | 1.28*** | 2.45*** | 2.25** | 2.07*** |
| | (.17) | (.9) | (.93) | (.61) |
| $\alpha_W$ | 0.60*** | 0.19*** | 0.11* | 0.15*** |
| | (.06) | (.06) | (.06) | (.05) |

NOTES: This table shows TSLS estimates for the labor mobility equation (33) when $\ln A^S_n = 0$ and $w^z_n = w^{CPS}_n$, i.e., there is no unobserved worker heterogeneity. The data are at the state-year level. Each column has 712 observations. The Cragg-Donald Wald F-statistic is 9.9 and the Kleibergen-Paap Wald F-statistic is 7.8 for the 2SLS specification in column (4). Every specification includes state and year fixed effects. Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

## D.7 Supplemental: 2SLS Estimates of Firm Parameters

This section presents both OLS and 2SLS estimates of the auxiliary parameters $b_0$, $b_1$, and $b_2$ in equation (40). When computing 2SLS estimates, we instrument for after-tax market potential, unit costs, and real government services using either the instrument vector of external tax rates $\boldsymbol{Z}^T_{nt} = (t^{*c}_{nt}, t^{*x}_{nt}, t^{*y}_{nt})$ and $MP^*_{nt}$, the vector of Bartik instruments $\boldsymbol{Z}^B_{nt} \equiv (\text{BtkP}_{nt}, \text{BtkTR}_{nt})$ and $MP^*_{nt}$, or all of these instruments combined.

As mentioned in the main text, our model predicts that OLS estimates of $b_0$, $b_1$, and $b_2$ are asymptotically biased due to the dependence of after-tax market potential, costs, and government spending in state $n$ and year $t$ on unobserved productivity or government efficiency in the same state and year, which are accounted for in the term $\nu^M_{nt}$.

Table A.13 provides the estimates of the first-stage regression corresponding to the 2SLS estimation of $b_0$, $b_1$, and $b_2$. The table shows how after-tax market potential, unit costs, and real government spending relate to the instruments. Column (1) shows the estimates of a regression of after-tax market potential on the instrument vector $\boldsymbol{Z}^T_{nt}$, the leave-out market potential term, and state and year fixed effects. Column (2) replaces with $\boldsymbol{Z}^B_{nt}$, and column (3) uses both instrument vectors. These three columns show that the leave-out market potential term is highly correlated with after-tax market potential. Columns (4)-(6) show similar specifications for unit costs, which tend to be lower when the state is close to high sales tax and low market potential neighbors. Columns (7)-(9) show similar results for real tax revenues, which tend to be high when leave-out market potential is high and when the that state's main tax revenue source is high nationally.

To increase power and mimic the variation used to estimate $\varepsilon_F$ in those cases in which we calibrate the value of $\alpha_F$, Table A.14 reports first-stage estimates for the combinations of after-tax market potential, unit costs, and real government spending used to identify $\varepsilon_F$ in these cases. Specifically, in the case in which we assume that $\alpha_F = 0.04$, we can write the right hand side of equation (40) as $b_0 \times RHS_{nt}$, where $RHS_{nt} \equiv \ln((1 - \bar{t}_{nt})MP_{nt}) - (\sigma - 1)\ln c_{nt} + 0.05(\sigma - 1)\ln \tilde{R}_{nt}$, and $\sigma$ is calibrated to equal 4. Similarly, in the case in which we assume that $\alpha_F = 0$, we can write the right-hand side of equation (40) as $b_0 \times RHS_{nt}$, where $RHS_{nt} \equiv \ln((1 - \bar{t}_{nt})MP_{nt}) - (\sigma - 1)\ln c_{nt}$. Columns (1)-(3) and (4)-(6) report the first stage estimates for $RHS_{nt}$ for these two possible calibrations of the parameter $\alpha_F$, respectively. The composite term tends to be positively correlated with nearby state tax rates and leave-out market potential.

Table A.13: First Stage of Firm-Location Equation

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | $\ln((1-\bar{t}_{nt})MP_{nt})$ | | | $\ln c_{nt}$ | | | $\ln \tilde{R}_{nt}$ | | |
| | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ |
| $t^{*x}_{nt}$ | 2.30** | | 2.34** | 0.31 | | 0.30 | -1.02 | | -1.10* |
| | (0.96) | | (0.96) | (0.22) | | (0.22) | (0.64) | | (0.63) |
| $t^{*y}_{nt}$ | 3.39** | | 3.26** | 0.17 | | 0.21 | -1.24 | | -0.94 |
| | (1.59) | | (1.62) | (0.41) | | (0.42) | (1.52) | | (1.45) |
| $t^{*c}_{nt}$ | 0.60 | | 0.60 | -1.32*** | | -1.31*** | 1.79 | | 1.85 |
| | (1.97) | | (1.98) | (0.46) | | (0.46) | (1.65) | | (1.65) |
| $\ln MP_{!n,t}$ | 2.72*** | 2.48*** | 2.72*** | 0.26*** | 0.28*** | 0.26*** | 0.90*** | 0.93*** | 0.91*** |
| | (0.40) | (0.40) | (0.40) | (0.08) | (0.08) | (0.08) | (0.26) | (0.26) | (0.26) |
| $BtkW_{nt}$ | | 0.01 | 0.01 | | 0.00 | 0.00 | | 0.00 | 0.00 |
| | | (0.01) | (0.01) | | (0.00) | (0.00) | | (0.01) | (0.01) |
| $BtkTR_{nt}$ | | -0.10 | -0.09 | | 0.03 | 0.03 | | 0.19 | 0.20 |
| | | (0.18) | (0.17) | | (0.06) | (0.06) | | (0.16) | (0.15) |
| R-squared | 0.996 | 0.996 | 0.996 | 0.993 | 0.993 | 0.993 | 0.995 | 0.995 | 0.995 |
| F-stat | 12.6 | 13.5 | 8.7 | 7.7 | 4.6 | 5.1 | 4.7 | 4.7 | 3.3 |

NOTES: This table shows first stage estimates for the firm mobility equation (40). The dependent variables are after-tax market potential in columns 1-3, unit cost in columns 4-6, and real government expenditures in columns 7-9. The data are at the state-year level. Every specification includes state and year fixed effects. Each row has 587 observations. Observations are weighted by state population. Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

Table A.14: First Stage of Firm-Location Equation

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | $RHS$ with $\alpha_F = .04$ | | | $RHS$ with $\alpha_F = 0$ | | |
| | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}$ | $\boldsymbol{Z}^B_{nt}$ | $\boldsymbol{Z}^T_{nt}, \boldsymbol{Z}^B_{nt}$ |
| $t^{*x}_{nt}$ | 1.25** | | 1.31** | 1.37*** | | 1.44*** |
| | (0.57) | | (0.57) | (0.53) | | (0.53) |
| $t^{*y}_{nt}$ | 2.74** | | 2.52** | 2.88*** | | 2.63** |
| | (1.11) | | (1.12) | (1.08) | | (1.08) |
| $t^{*c}_{nt}$ | 4.79*** | | 4.76*** | 4.57*** | | 4.53*** |
| | (1.52) | | (1.53) | (1.41) | | (1.43) |
| $\ln MP_{!n,t}$ | 2.06*** | 1.76*** | 2.06*** | 1.95*** | 1.65*** | 1.95*** |
| | (0.29) | (0.32) | (0.29) | (0.27) | (0.30) | (0.27) |
| $BtkW_{nt}$ | | 0.00 | 0.01 | | 0.00 | 0.01 |
| | | (0.01) | (0.01) | | (0.01) | (0.01) |
| $BtkTR_{nt}$ | | -0.16 | -0.15 | | -0.19 | -0.17 |
| | | (0.16) | (0.15) | | (0.15) | (0.14) |
| R-squared | 0.996 | 0.995 | 0.996 | 0.995 | 0.995 | 0.995 |
| F-stat | 16.2 | 11.2 | 11.1 | 17.3 | 11.5 | 12 |

NOTES: This table shows first stage estimates for the firm mobility equation (40). The dependent variables are two versions of the variable $RHS = \ln((1-\bar{t}_{nt})MP_{nt}) - (\sigma-1)\ln c_{nt} + \alpha_F(\sigma-1)\ln \tilde{R}_{nt}$. Columns 1-3 show estimates for the sum of after-tax market potential, $(\sigma-1)=3$ times unit costs, and $\alpha_F \times (\sigma-1) = .04 \times 3$ times real government expenditures (which results in common coefficients in the model). Similiarly, columns 4-6 correspond to columns 1-3 with $\alpha_F = 0$, so the sum is just of after-tax market potential and 3 times unit costs. Observations are weighted by state population. Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

Table A.15 presents OLS and 2SLS estimates of $b_0$, $b_1$, and $b_2$. Columns (1)-(3) present OLS estimates and (4)-(12) present 2SLS estimates. Column (1) shows that higher after-tax market potential and real government services tend to attract firms and that higher costs are unattractive. Recall that $(\varepsilon_F, \alpha_F)$ are over-identified, but that the ratio of $-b_2/b_1$ identifies $\alpha_F$. Intuitively, firm location is $0.42/0.14 = 3$ times as responsive to unit costs as to real government spending, and $\alpha_F = 1/3 = .34$ reflects the inverse of this relative responsiveness. Columns (2) and (3) show the OLS estimate of $b_0$ in the cases in which we either assume that $\alpha_F$ is equal to the cross-state average $R_n/GDP_n$ or we set it to 0; the resulting estimate of $b_0$ is similar to that in column (1). Our model predicts that these OLS estimates are asymptotically biased estimates of the parameters $b_0$, $b_1$, and $b_2$, the reason being that after-tax market potential, production costs and real government services are likely correlated with unobserved state productivity and government efficiency.

Column (4) in Table A.15 shows that the 2SLS estimates are larger than the OLS estimates for the coefficients on after-tax market potential and real government services and smaller than the corresponding OLS estimate for the coefficient on unit costs. The coefficient on real government services is estimated imprecisely: this shows that the identification of the structural parameters $\varepsilon_F$ and $\alpha_F$ in our GMM estimation approach comes mainly from the auxiliary parameters $b_0$ and $b_1$. Furthermore, as columns (5) and (6) illustrate, conditional on calibrated values of $\alpha_F$, the 2SLS estimate of parameter $\varepsilon_F$ is estimated with a high degree of precision. Specifically, columns (11)-(12) show an estimate of 0.7 for the 2SLS estimate of the parameter $b_0$. Given that $b_0 \equiv (\varepsilon_F / (\sigma - 1)) / (1 + \chi_F \alpha_F \varepsilon_F)$, an estimate of 0.7 for $b_0$ implies that $\hat{\varepsilon}_F = ((\sigma - 1)(\hat{b}_0))/(1 - \chi_F \alpha_F (\sigma - 1)) = (3 \times .7)/(1 - .7 \times .04 \times 3) = 2.29$. This estimate of $\hat{\varepsilon}_F = 2.29$ is precise. Similarly, the 2SLS estimate of $\hat{\varepsilon}_F$ under the assumption that $\alpha_F = 0$ is also precisely estimated. Moreover, the estimates in columns (11) and (12) are not affected by weak instrument problems. The Cragg-Donald Wald F-statistic is 16.7 and the Kleibergen-Paap Wald F-statistic is 11.1 for the 2SLS specification in column (11) and 17.6 and 12.0, respectively, for the specification in column (12).

Table A.15: OLS and 2SLS Estimates of Firm-Location Parameters

| | (1) OLS | (2) OLS | (3) OLS | (4) 2SLS | (5) 2SLS | (6) 2SLS | (7) 2SLS | (8) 2SLS | (9) 2SLS | (10) 2SLS | (11) 2SLS | (12) 2SLS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $\boldsymbol{Z}_{nt}^{T}$ | $\boldsymbol{Z}_{nt}^{T}$ | $\boldsymbol{Z}_{nt}^{T}$ | $\boldsymbol{Z}_{nt}^{B}$ | $\boldsymbol{Z}_{nt}^{B}$ | $\boldsymbol{Z}_{nt}^{B}$ | $\boldsymbol{Z}_{nt}^{T}, \boldsymbol{Z}_{nt}^{B}$ | $\boldsymbol{Z}_{nt}^{T}, \boldsymbol{Z}_{nt}^{B}$ | $\boldsymbol{Z}_{nt}^{T}, \boldsymbol{Z}_{nt}^{B}$ |
| $\ln(1-\bar{t}_{nt})MP_{nt}$ | 0.34*** | | | 0.78*** | | | 0.25 | | | 0.71*** | | |
| | (0.04) | | | (0.14) | | | (0.82) | | | (0.13) | | |
| $\ln c_{nt}$ | -0.42*** | | | -3.00*** | | | 4.80 | | | -2.64*** | | |
| | (0.11) | | | (0.72) | | | (7.64) | | | (0.71) | | |
| $\ln \tilde{R}_{nt}$ | 0.14*** | | | 0.01 | | | -0.91 | | | 0.13 | | |
| | (0.04) | | | (0.26) | | | (1.82) | | | (0.22) | | |
| $RHS$ with $\alpha_F = .04$ | | 0.39*** | | | 0.69*** | | | 0.62*** | | | 0.68*** | |
| | | (0.03) | | | (0.07) | | | (0.08) | | | (0.07) | |
| $RHS$ with $\alpha_F = 0$ | | | 0.40*** | | | 0.72*** | | | 0.65*** | | | 0.70*** |
| | | | (0.03) | | | (0.07) | | | (0.09) | | | (0.08) |
| Structural Parameters | | | | | | | | | | | | |
| $\varepsilon_F$ for $\chi_F = 0$ | 1.03*** | | | 2.33*** | | | .75 | | | 2.12*** | | |
| | (.11) | | | (.41) | | | (2.46) | | | (.38) | | |
| $\varepsilon_F$ for $\chi_F = 1$ | 1.59*** | | | 2.34*** | | | .88 | | | 2.28*** | | |
| | (.26) | | | (.33) | | | (3.58) | | | (.29) | | |
| $\alpha_F$ | 0.34** | | | 0.00 | | | 0.19 | | | .05 | | |
| | (.17) | | | (.09) | | | (.35) | | | (.09) | | |

NOTES: This table shows OLS and 2SLS estimates. The dependent variable in each column is log of the number of establishments $\ln M_{nt}$. The data are at the state-year level. Each column has 587 observations. The dependent variables are after-tax market potential, unit cost, and real government expenditures. $RHS$ is $\ln((1-\bar{t}_{nt})MP_{nt}) - (\sigma-1)\ln c_{nt} + \alpha_F(\sigma-1)\ln \tilde{R}_{nt}$. Every specification includes state and year fixed effects. Robust standard errors are in parentheses and *** p<0.01, ** p<0.05, * p<0.1.

## D.8 Supplemental: Dynamic Panel Data Elasticities

The model described in Section 4 is a static model, and thus assumes that workers and firms can move across locations freely, without any need to pay a fixed costs of moving. Consequently, the equilibrium equations used to estimate the structural elasticities of labor and firm mobility with respect to changes in taxes, economic variables and public spending (i.e. (33) and (40) in the main text) predict that the share of workers and firms located in each state in any given year $t$ depends exclusively on the period $t$ values of different covariates. In a more general model with fixed costs of mobility, the population or firm share in a location in a period $t$ will depend on the corresponding share in every location in period $t-1$. Furthermore, in this general model, a permanent change in any of the economic determinants of workers' and firms' locations in a period $t$ will have a different impact on the short run (i.e. in the same period $t$) and on the long run (i.e., infinite periods ahead).

In this Appendix section, we explore how the static panel data elasticities that we estimate following the procedure in Section 6.2 compare to the short-run and long-run elasticities generated by a dynamic panel data model with multiple locations. Specifically, for a set of locations $i = 1, \ldots, 50$ and time periods $t = 1, \ldots, 1000$, we simulate the following statistical model:

$$l_{it} = \beta\rho_l l_{it-1} + (1-\beta)\rho_l(N-1)^{-1}\sum_{n\neq i} l_{nt-1} + x_{it} + \varepsilon_{l,it} \tag{A.59}$$

$$x_{it} = \alpha_{0,i} + \rho_x x_{it-1} + \varepsilon_{x,it} \tag{A.60}$$

$$\varepsilon_{l,it} \sim \mathbb{N}(0,1) \quad \text{and independent across } i \text{ and } t,$$

$$\varepsilon_{x,it} \sim \mathbb{N}(0,1) \quad \text{and independent across } i \text{ and } t, \tag{A.61}$$

$$\alpha_{0,i} \sim \mathbb{N}(0,1) \quad \text{and independent across } i, \tag{A.62}$$

$$(l_{i0}, x_{i0}) = (0,0). \tag{A.63}$$

According to (A.59), (log) population (or firms or workers) in a location $i$ in a period $t$, $l_{it}$, depends both on the population in every state in period $t-1$, $\{l_{it-1}\}_{i=1}^{50}$, and on the regressor $x_{it}$. The coefficient on $x_{it}$ is assumed to be equal to one. As reflected in (A.59), the parameter vector $\rho_l$ indicates the degree to which the (log) population in any location $i$ is affected by the distribution of population across locations in period $t-1$. Specifically, if $\rho_l = 0$, then(A.59) is static and, thus, there is no serial correlation in population. The opposite is true if $\rho_l$ is close to one. Given a value of $\rho_l$, the parameter $\beta$ indicates the degree to which population in a location $i$ is affected by past population in the same location $i$. If $\beta = 1$, equation (A.59) implies there is no migration across regions. The opposite is true when $\beta$ is equal to zero.

Equation (A.60) indicates the time evolution of $x_{it}$. Specifically, the parameter $\rho_x$ modulates the degree of persistence in $x_{it}$. The variable $x_{it}$ plays the role of after-tax real wages or real government spending in the labor mobility equation in (33), and the role of market potential, unit cost or real government spending in the firm mobility equation in (40).

For each combination of the following parameter values

$$\rho_l \in \{0, 0.1, 0.5, 0.9\}, \tag{A.64}$$

$$\rho_x \in \{0.5, 0.9\}, \tag{A.65}$$

$$\beta \in \{0.5, 0.9\}, \tag{A.66}$$

we simulate 1000 different longitudinal datasets using (A.59) to (A.63).

For each of the 1000 simulated datasets corresponding to a particular parameter vector $(\rho_l, \rho_x, \beta)$, we form an estimation sample by keeping the information on the simulated values of $l_{it}$ and $x_{it}$ for the last 25 periods we simulate and for all the 50 locations in our simulated dataset. Our choice of the number of periods and locations in the simulated estimation sample aims to replicate the dimensions of the sample that we use for estimation in Section 6.2. By keeping only the last 25 periods of the simulated dataset, we make sure that the observations that we keep in our simulated estimation sample are unaffected by the initial conditions $(l_{i0}, x_{i0})$.

For each of the 1000 generated estimation samples corresponding to a particular parameter vector $(\rho_l, \rho_x, \beta)$, we use Ordinary Least Squares (OLS) to estimate a static linear panel data model analogous to that in(33) and (40):

$$l_{it} = \gamma_i + \gamma_t + \gamma_l x_{it} + u_{it},$$

where $\gamma_i$ denotes a location $i$ fixed effect, $\gamma_t$ denotes a period $t$ fixed effect, and $u_{it}$ is an unobserved residual.

For each parameter vector $(\rho_l, \rho_x, \beta)$ that we explore in our simulation, Table A.16 reports: (a) the mean and standard deviation of the OLS estimates $\hat{\gamma}_l$ that we obtain in our 1000 generated estimation samples; (b) the true short-run and long-run impact on the dependent variable $l_i$ for a permanent change in one unit in $x_i$.

Table A.16: Elasticity Simulation Estimates

| | $\rho_l = 0$ | | | | $\rho_l = 0.1$ | | | |
|---|---|---|---|---|---|---|---|---|
| | $\rho_x = 0.5$ | | $\rho_x = 0.9$ | | $\rho_x = 0.5$ | | $\rho_x = 0.9$ | |
| | $\beta = 0.5$ | $\beta = 0.9$ | $\beta = 0.5$ | $\beta = 0.9$ | $\beta = 0.5$ | $\beta = 0.9$ | $\beta = 0.5$ | $\beta = 0.9$ |
| $\bar{\hat{\gamma}}_l$ | 1 | 1 | 1 | 1 | 1.02 | 1.04 | 1.04 | 1.08 |
| $s.d.(\hat{\gamma}_l)$ | (0.03) | (0.05) | (0.03) | (0.03) | (0.03) | (0.03) | (0.02) | (0.02) |
| Long-Run Impact | 1 | 1 | 1 | 1 | 1.05 | 1.10 | 1.05 | 1.10 |
| Short-Run Impact | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $\rho_l = 0.5$ | | | | $\rho_l = 0.9$ | | | |
| | $\rho_x = 0.5$ | | $\rho_x = 0.9$ | | $\rho_x = 0.5$ | | $\rho_x = 0.9$ | |
| | $\beta = 0.5$ | $\beta = 0.9$ | $\beta = 0.5$ | $\beta = 0.9$ | $\beta = 0.5$ | $\beta = 0.9$ | $\beta = 0.5$ | $\beta = 0.9$ |
| $\bar{\hat{\gamma}}_l$ | 1.13 | 1.25 | 1.27 | 1.62 | 1.25 | 1.48 | 1.65 | 3.23 |
| $s.d.(\hat{\gamma}_l)$ | (0.03) | (0.05) | (0.03) | (0.07) | (0.08) | (0.23) | (0.35) | (0.66) |
| Long-Run Impact | 1.34 | 1.82 | 1.34 | 1.82 | 1.95 | 5.31 | 1.95 | 5.31 |
| Short-Run Impact | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

The results in Table A.16 illustrate that, no matter the value of the parameters $\rho_l$, $\rho_x$, and $\beta$, our estimates tend to be between the short-run and the long-run impact parameters. Furthermore, the quantitative difference between our point estimates and the true long-run impact increases in the value of the three parameters, reaching its maximum when $\rho_l = \rho_x = \beta = 0.9$.

The OLS estimate of the coefficient of each of the regressors entering either the labor mobility or the firm mobility equations on their own respective lag is always close to 0.9. Therefore, the relevant value of $\rho_x$ is close to 0.9. It is reasonable to expect that the population of a state in a year $t$ depends significantly more on the lag population of the same state than on the lag population in other states, so the actual value of $\beta$ is probably larger or equal than 0.5. Similarly, the actual value of the parameter $\rho_l$ is also likely larger or equal than 0.5, reflecting a significant amount of persistence in each state's population. Looking at the relevant cells of Table \ref{tab: simul}, one can conclude that, given the value of $\rho_x$ close to 0.9: (a) if either $\beta$ or $\rho_l$ are close to 0.5, then the estimate $\hat{\gamma}_l$ will likely be very close to the long-run impact of the regressor on the dependent variable; (b) only if both $\beta$ and $\rho_l$ are very close to 0.9, the estimate $\hat{\gamma}_l$ will likely be half-way between the short-run and the long-run impact of the regressor on the dependent variable.

## D.9 Comparison with Existing Estimates

Researchers have previously estimated regressions similar to (33) and (40) using sources of variation different from ours to identify the labor and firm mobility elasticities. Table A.17 compares our estimates of $\varepsilon_W$, $\alpha_W$, $\varepsilon_F$, and $\alpha_F$ to those that we would have constructed if we had used estimates of the elasticity of labor and firms with respect to after-tax wages and public expenditure from six recent studies. The parameter that is most often estimated is the elasticity of labor with respect to real wages; this previous literature implies estimates of $\varepsilon_W$ with mean value of 1.81. Our numbers of $\varepsilon_W = 1.36$ ($\chi_W = 0$) and $\varepsilon_W = 1.73$ ($\chi_W = 1$) reported in Table 1 are within the range of these estimates. Our estimate of $\varepsilon_F$ is between the firm-mobility parameters reported in Suárez Serrato and Zidar (2016) and Giroud and Rauh (2015).

Concerning $\alpha_W$ and $\alpha_F$, there is substantial evidence that public expenditures have amenity and productivity value for workers and firms, respectively, which is consistent with $\alpha_W > 0$ and $\alpha_F > 0$. Some studies infer positive amenity value for government spending from land rents,[5] while others focus on the productivity effects of large investment projects.[6] However, very few papers estimate specifications similar to (33) and (40). The estimates of the effects of variation in federal spending at the local level from Suárez Serrato and Wingender (2016) imply $\alpha_F = 0.10$ and $\alpha_W = 0.26$.

Of course, all these comparisons are imperfect due to differences in the source of variation, geography, and time dimension; for example, all of these studies use smaller geographic units than states. Additionally, not all specifications include the same covariates as in (33) and (40). These differences notwithstanding, our structural parameters are close to those in the literature.

[5]E.g., Bradbury et al. (2001) show that local areas in Massachusetts with lower increases in government spending had lower house prices, and Cellini et al. (2010) show that public infrastructure spending on school facilities raised local housing values in California. Their estimates imply a willingness to pay \$1.50 or more for each dollar of capital spending. Chay and Greenstone (2005) and Black (1999) also provide evidence of amenity value from government regulations on air quality and from school quality, respectively.

[6]Kline and Moretti (2014) find that infrastructure investments in by the Tennessee Valley Authority resulted in large and direct productivity increases, yielding benefits that exceeded the costs of the program. Fernald (1999) also provides evidence that road-building increases productivity, especially in vehicle-intensive industries. Haughwout (2002) shows evidence from a large sample of U.S. cities that "public capital provides significant productivity and consumption benefits" for both firms and workers.

Table A.17: Structural Parameters Implied by Similar Studies

| Paper | Estimates | Implied Values of | | | | Source of Variation (Shock) | Level of Variation |
|---|---|---|---|---|---|---|---|
| | | $\varepsilon_W$ | $\alpha_W$ | $\varepsilon_F$ | $\alpha_F$ | | |
| Bound and Holzer (2000) | $a_0 = 1.20$[7] | 1.16 | | | | Bartik | MSA (1980's) |
| Notowidigdo (2013) | $a_0 = 3.47$[8] | 2.49 | | | | Bartik | MSA (1980-2000) |
| Suárez Serrato and Wingender (2016) | $a_0 = 1.58$[9] | 1.45 | | | | Bartik and Census Instrument | County Group (1980-2009) |
| | $a_0 = 2.9$, $a_1 = 1.02$, $b_1 = 0.26$[10] | 1.94 | 0.26 | | 0.10 | | |
| Diamond (2016) | $a_0 = 3.10$[11] | 2.32 | | | | Bartik | MSA (1980-2000) |
| Suárez Serrato and Zidar (2016) | $a_0 = 1.28$[12] | 1.23 | | | | Bartik | County Group (1980-2009) |
| | $a_0 = 2.63$, $b_0 = 3.35$[13] | 2.09 | | 5.26 | | Business Tax | |
| Giroud and Rauh (2015) | $b_0 = 0.40$[14] | | | 1.31 | | Corporate Tax | Firm-Level (1977-2011) |

NOTES: This table reports the values of our structural parameters implied by estimates of specifications similar to (33) and (40) found in the previous literature. Whenever needed, we assume the values used in our baseline parametrization of $\sigma = 4$, $\chi_W = 1$, $\chi_F = 1$, $\alpha_F = 0.03$, and $\alpha_W = 0.16$ in recovering structural parameters. When the effects are only reported separately for skilled and unskilled workers we use a share of skilled workers of 33% to average the effects.



[7] For both college and non-college groups, we first construct $a_0$ from Table 3 in Bound and Holzer (2000) by taking the ratio of the effects on Population and Total Hours. We then average the effect by the college share above.

[8] This parameter comes from Table 3 in Notowidigdo (2013) and results from taking the ratio of columns (1) and (6). Note that these specifications also control for quadratic effects. We employ marginal effects around 0.

[9] This number is directly reported in Suárez Serrato and Wingender (2016) in Table 9.

[10] The parameters $a_0$ and $a_1$ come from Table 10 in Suárez Serrato and Wingender (2016) by manipulating the structural parameters as follows: $a_0 = 1/\sigma^i$ and $a_0 = \psi^i/\sigma^i$ for each skill group. The parameter $b_1$ comes from using the effect of spending on firm location and by noting that this effect is equal to $1 - (\kappa_i^{GS} + (1-\kappa_i^{GS})/(1-\alpha_i))\frac{\partial W^i}{\partial F}$ in Suárez Serrato and Wingender (2016). The parameters $\alpha^i$,$\kappa_i^{GS}$, and $\frac{\partial W^i}{\partial F}$ are reported in Tables 9 and 10 by skill group in Suárez Serrato and Wingender (2016). We then average these effects by the college share above.

[11] Diamond (2016) reports the effect on wage on population by skill group in Table 3. We then average these effects by the college share above. Note that Diamond (2016) also controls for state of origin which leads to a larger effect of population on wages than in other similar papers, especially for the low skill population.

[12] We construct $a_0$ from Table 6, Panel (c) in Suárez Serrato and Zidar (2016) by taking the ratio of the effects on Population and Wages.

[13] We construct $a_0$ from Table 6, Panel (c) in Suárez Serrato and Zidar (2016) by taking the ratio of the effects on Population and Wages. $b_0$ is reported in Table 6, Panel (c).

[14] Giroud and Rauh (2015) report an elasticity of number of establishment with respect to corporate taxes of 0.4.

# E Appendix to Section 7

## E.1 Consumption-Equivalent Welfare Change

The change in indirect utility $\hat{v}$ in (45) follows from assuming that an individual $l$ located in $n$ receives utility $v_n \epsilon_n^l$. Assume that, instead of $v_n \epsilon_n^l$, the indirect utility received by an individual $l$ in $n$ was $W\left(v_n \epsilon_n^l\right)$ where $W(\cdot)$ is an increasing function. We continue to assume that the $\epsilon_n^l$ are i.i.d normalized Fréchet, with CDF $H(x) = \Pr\left(\epsilon_n^l < x\right) = e^{-x^{-\varepsilon}}$. This monotone transformation does not impact the choice probabilities. As we show below, in this case the expected indirect utility of a worker is

$$V = \mathbb{E}\left[W(v\epsilon)\right] \equiv \int_{\epsilon} W(v\epsilon)\, dH(\epsilon), \tag{A.67}$$

where $v$ is defined in (8) and the expectation is over the Fréchet draw $\epsilon$. Therefore, in any counterfactual the relative change in indirect utility is

$$\hat{V} = \frac{\mathbb{E}\left[W(v\hat{v}\epsilon)\right]}{\mathbb{E}\left[W(v\epsilon)\right]}, \tag{A.68}$$

implying that any change in private or public consumption that leads to a relative change $\hat{v}$ is equivalent to the actual welfare change experienced by workers under any monotone function $W(\cdot)$. For example, the counterfactual change in welfare $\hat{v}$ is equivalent to the change in welfare that would arise if the relative consumption of both the public and the private goods were to change in every state by an amount equal to $\hat{v}$.

To derive (A.67), by definition of $V$ we have:

$$V = \sum_n \int_{\epsilon} \Pr\left[v_{n'}\epsilon_{n'} \leq v_n \epsilon, \forall n' \neq n\right] W(v_n\epsilon)\, H'(\epsilon)\, d\epsilon. \tag{A.69}$$

Since the shocks are i.i.d, we have $\Pr\left[v_{n'}\epsilon_{n'} \leq v_n\epsilon, \forall n' \neq n\right] = \prod_{n' \neq n} H\left(c_n \epsilon / c_{n'}\right)$. Using this expression, the definition of $H(x)$ and the change of variable $z_n \equiv v_n \epsilon$ we get:

$$V = \sum_n \int_{z_n} \prod_{\forall n'} H\left(\frac{z_n}{v_{n'}}\right) W(z_n)\, \varepsilon \left(\frac{z_n}{v_n}\right)^{-\varepsilon-1} \frac{1}{v_n} dz_n. \tag{A.70}$$

We also have that $\prod_{\forall n'} H\left(z_n / c_{n'}\right) = H\left(z_n / v\right)$ for $v$ defined in (8). Using this property and the change of variable $Z_n = z_n / v$ we get

$$V = \sum_n \int_{Z_n} W(vZ_n)\, \varepsilon \left(\frac{Z_n}{v_n} v\right)^{-\varepsilon-1} \frac{v}{v_n} H(Z_n)\, dZ_n, \tag{A.71}$$

which, using the definition of $H'(x)$, gives:

$$V = \sum_n \left(\frac{v_n}{v}\right)^{\varepsilon} \int_{Z_n} W(vZ_n)\, dH(Z_n). \tag{A.72}$$

Using the fact that $\sum_n (v_n/v)^{\varepsilon} = 1$ gives (A.67).

## E.2 Derivation of Equation (46)

It follows from (7) and (8) that, in any counterfactual, the first-order approximation to the change in $v$ is:

$$d\ln v = \sum L_n d\ln v_n. \tag{A.73}$$

In what follows, we consider a special case with no trade costs ($\tau_{in} = 1$ for all $i, n$), perfect substitutability across varieties ($\sigma \to \infty$), homogeneous firms ($\varepsilon_F \to \infty$), constant labor supply ($h_n$ constant), and identical preferences for government spending across states ($\alpha_{W,n} = \alpha_W$). In addition, we also consider a tax structure with only state sales and income taxes. These assumptions are more general than the restrictions in Proposition 1, in that we do not impose non-rival public goods and we allow for dispersion in preference draws, $\varepsilon_W < \infty$, and in individual productivity draws, $\zeta_n < \infty$. In addition, we assume that state income and sales taxes are the only taxes, and that income taxes are constant. As a result, the tax distribution is characterized by the keep rate $1-T_n = (1-t_n^y)/(1+t_n^c)$ of each state.

Then, because we have assumed constant intensive margin of labor supply, we have:

$$d\ln v_n = (1-\alpha_W)\, d\ln\left(\frac{C_n^L}{L_n}\right) + \alpha_W d\ln\frac{G_n}{L_n} + \alpha_W\left(1-\chi_W\right) d\ln L_n, \tag{A.74}$$

where $C_n^L$ is the aggregate consumption of workers. Because firms make zero profits, (A.8) implies that the payments to capital owners are a fraction $\beta$ of the GDP of state $n$, $GDP_n$. Therefore, the final consumption of capital owners is $C_n^K = (1-T_n)\,\beta GDP_n$. In addition, under our assumptions, government spending is $G_n = T_n GDP_n$. Combining these last two expressions with the fact that $C_n^W + C_n^K + G_n = GDP_n$, aggregate consumption of workers can be written: $C_n^L = (1-\beta)\left(GDP_n - G_n\right)$. This las expression in turn implies:

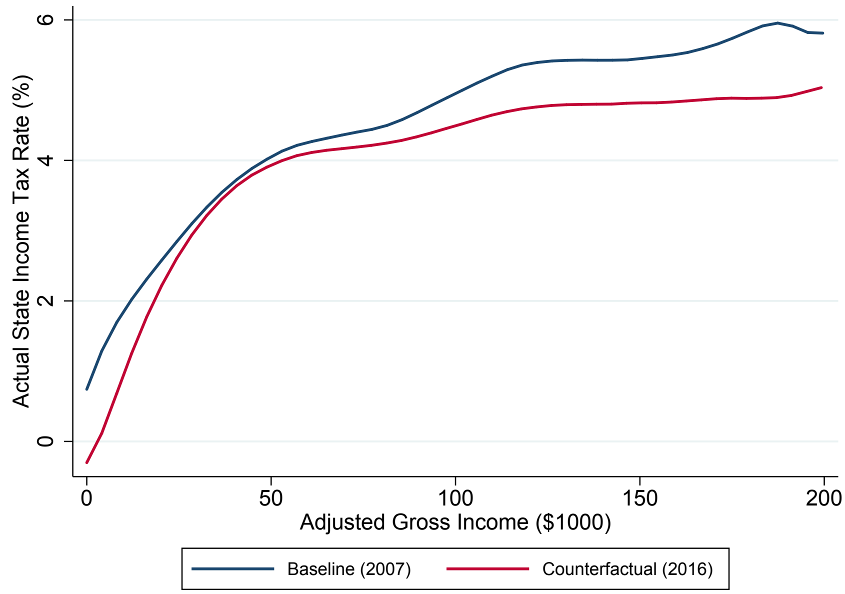
$$d\ln\frac{C_n^L}{L_n} = \frac{d\ln\frac{GDP_n}{L_n} - \frac{G_n}{GDP_n} d\ln\frac{G_n}{L_n}}{1-\frac{G_n}{GDP_n}}. \tag{A.75}$$

Combining (A.73), (A.74) and (A.75), assuming $\alpha_{W,n} = \alpha_W$ and from the labor market clearing condition that $\sum L_n dL_n = 0$ gives (46).

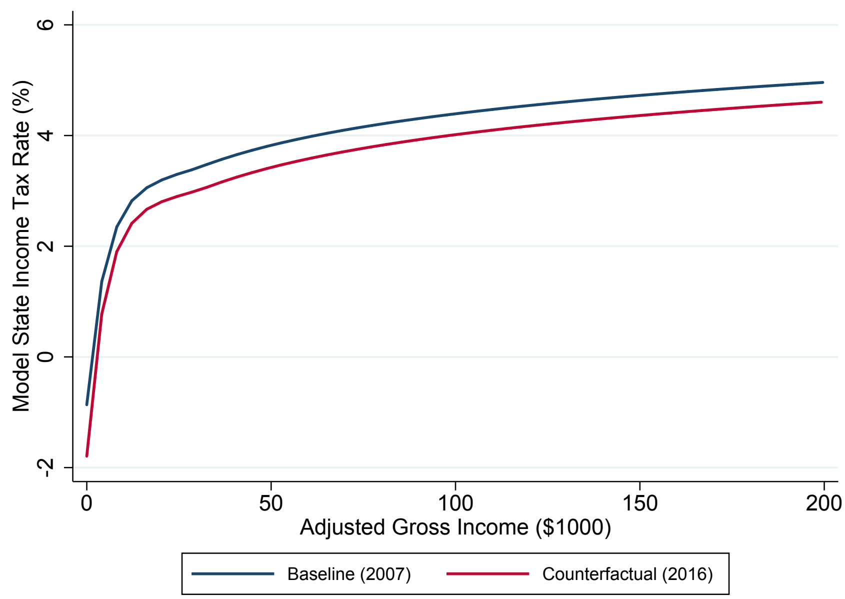
## E.3 Appendix Figures to Section 7.1

Figure A.3: The North Carolina Income Tax Cuts: Actual and Estimated Income Tax Schedules Before and After the Reform

(a) Actual Income Tax Rates



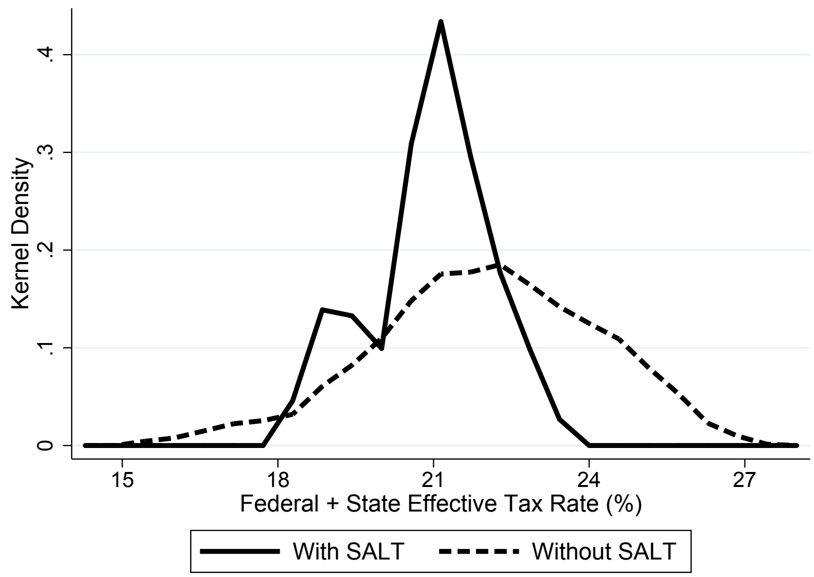

(b) Model Income Tax Rates




Notes: This figure plots actual and model-based income tax rates in North Carolina before and after the 2014-2016 reform.

## E.4 Appendix to Section 7.3

Figure A.4: Dispersion in Federal and State Income Taxes with and without SALT




Notes: This figure shows dispersion in effective tax rates when State and Local Taxes (SALT) can or cannot be deducted from federal income. Tax rates are computed using a sample of individual tax returns from the Statistics of Income (SOI) Public Use Files and NBER's tax simulator TAXSIM. Individual returns with negative Adjusted Gross Income (AGI) are dropped and the remaining observations are winsorized by their effective federal and state income tax rate at the 1st and 99th percentiles. The figure displays the kernel density of federal and state income tax rates in 2007.

Table A.18: Income Tax Parameters and Effective Rates with and without SALT

| State | With SALT | | Without SALT | | Rates with SALT if AGI is | | | | Rates without SALT if AGI is | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $a_n^y$ | $b_n^y$ | $a_n^y$ | $b_n^y$ | 25K | 50K | 100K | 200K | 25K | 50K | 100K | 200K |
| AL | 1.273 | 0.044 | 1.218 | 0.039 | 14.8 | 18.1 | 22.9 | 25.2 | 14.7 | 17.6 | 21.9 | 24.0 |
| AK | 1.243 | 0.039 | 1.368 | 0.049 | 13.1 | 16.2 | 20.6 | 22.7 | 12.7 | 16.5 | 22.0 | 24.6 |
| AZ | 1.336 | 0.047 | 1.537 | 0.061 | 13.4 | 17.0 | 22.2 | 24.7 | 12.0 | 16.8 | 23.5 | 26.6 |
| AR | 1.352 | 0.049 | 1.437 | 0.053 | 14.1 | 17.9 | 23.3 | 25.9 | 12.2 | 16.4 | 22.3 | 25.1 |
| CA | 1.365 | 0.049 | 1.522 | 0.060 | 13.4 | 17.2 | 22.7 | 25.3 | 12.7 | 17.4 | 24.0 | 27.1 |
| CO | 1.322 | 0.047 | 1.372 | 0.053 | 14.2 | 17.8 | 23.0 | 25.5 | 15.6 | 19.6 | 25.3 | 27.9 |
| CT | 1.346 | 0.049 | 1.425 | 0.056 | 13.9 | 17.7 | 23.0 | 25.6 | 15.2 | 19.4 | 25.5 | 28.3 |
| DE | 1.323 | 0.047 | 1.487 | 0.058 | 14.0 | 17.6 | 22.8 | 25.2 | 12.7 | 17.2 | 23.6 | 26.6 |
| FL | 1.243 | 0.039 | 1.209 | 0.035 | 13.1 | 16.2 | 20.6 | 22.7 | 12.7 | 15.5 | 19.5 | 21.5 |
| GA | 1.400 | 0.053 | 1.859 | 0.079 | 13.8 | 17.9 | 23.7 | 26.4 | 10.6 | 16.8 | 25.5 | 29.5 |
| HI | 1.405 | 0.053 | 1.676 | 0.072 | 14.1 | 18.2 | 24.1 | 26.8 | 13.7 | 19.2 | 26.9 | 30.5 |
| ID | 1.440 | 0.055 | 1.648 | 0.068 | 13.6 | 17.9 | 24.0 | 26.8 | 12.1 | 17.5 | 24.9 | 28.4 |
| IL | 1.266 | 0.042 | 1.250 | 0.043 | 14.3 | 17.6 | 22.3 | 24.6 | 15.5 | 18.7 | 23.4 | 25.6 |
| IN | 1.266 | 0.043 | 1.332 | 0.049 | 15.0 | 18.3 | 23.1 | 25.3 | 15.5 | 19.3 | 24.6 | 27.2 |
| IA | 1.388 | 0.052 | 1.493 | 0.061 | 14.2 | 18.2 | 23.9 | 26.6 | 14.5 | 19.1 | 25.6 | 28.7 |
| KS | 1.322 | 0.047 | 1.238 | 0.043 | 14.5 | 18.1 | 23.3 | 25.8 | 17.0 | 20.3 | 24.9 | 27.1 |
| KY | 1.327 | 0.048 | 1.382 | 0.053 | 14.7 | 18.4 | 23.6 | 26.1 | 14.9 | 18.9 | 24.6 | 27.3 |
| LA | 1.341 | 0.048 | 1.794 | 0.075 | 14.0 | 17.7 | 23.0 | 25.5 | 10.4 | 16.4 | 24.7 | 28.5 |
| ME | 1.400 | 0.053 | 1.507 | 0.062 | 14.1 | 18.2 | 24.0 | 26.8 | 15.1 | 19.8 | 26.4 | 29.5 |
| MD | 1.308 | 0.046 | 1.272 | 0.046 | 14.4 | 17.9 | 23.0 | 25.4 | 17.0 | 20.5 | 25.4 | 27.8 |
| MA | 1.308 | 0.047 | 1.212 | 0.044 | 15.1 | 18.7 | 23.8 | 26.2 | 18.8 | 22.0 | 26.5 | 28.7 |
| MI | 1.302 | 0.046 | 1.468 | 0.057 | 14.4 | 17.9 | 22.9 | 25.3 | 13.3 | 17.8 | 24.0 | 27.0 |
| MN | 1.372 | 0.051 | 1.392 | 0.056 | 14.3 | 18.2 | 23.8 | 26.5 | 16.6 | 20.8 | 26.6 | 29.4 |
| MS | 1.255 | 0.041 | 1.052 | 0.022 | 14.3 | 17.5 | 22.1 | 24.3 | 13.9 | 15.6 | 18.2 | 19.4 |
| MO | 1.321 | 0.047 | 1.384 | 0.052 | 14.2 | 17.8 | 23.0 | 25.4 | 13.9 | 17.9 | 23.6 | 26.3 |
| MT | 1.354 | 0.049 | 1.440 | 0.054 | 14.1 | 17.9 | 23.3 | 25.9 | 12.5 | 16.7 | 22.7 | 25.6 |
| NE | 1.373 | 0.051 | 1.380 | 0.054 | 13.8 | 17.7 | 23.3 | 25.9 | 16.0 | 20.0 | 25.8 | 28.5 |
| NV | 1.243 | 0.039 | 1.371 | 0.049 | 13.1 | 16.2 | 20.6 | 22.7 | 12.7 | 16.5 | 22.0 | 24.6 |
| NH | 1.244 | 0.039 | 1.076 | 0.028 | 13.2 | 16.3 | 20.7 | 22.8 | 17.2 | 19.3 | 22.4 | 23.9 |
| NJ | 1.307 | 0.045 | 1.214 | 0.041 | 13.8 | 17.3 | 22.4 | 24.8 | 17.0 | 20.1 | 24.5 | 26.7 |
| NM | 1.462 | 0.056 | 2.053 | 0.088 | 12.5 | 16.8 | 23.0 | 25.9 | 8.4 | 15.5 | 25.1 | 29.6 |
| NY | 1.361 | 0.050 | 1.361 | 0.054 | 14.3 | 18.1 | 23.6 | 26.2 | 17.1 | 21.1 | 26.7 | 29.4 |
| NC | 1.309 | 0.047 | 1.384 | 0.053 | 15.2 | 18.8 | 23.9 | 26.4 | 14.7 | 18.7 | 24.4 | 27.1 |
| ND | 1.305 | 0.045 | 1.349 | 0.048 | 13.5 | 17.0 | 22.0 | 24.4 | 13.5 | 17.3 | 22.6 | 25.2 |
| OH | 1.316 | 0.046 | 1.317 | 0.048 | 14.1 | 17.7 | 22.8 | 25.2 | 15.2 | 18.8 | 24.0 | 26.5 |
| OK | 1.418 | 0.054 | 1.625 | 0.069 | 13.7 | 17.9 | 23.8 | 26.6 | 13.6 | 18.8 | 26.2 | 29.6 |
| OR | 1.370 | 0.052 | 1.449 | 0.059 | 15.4 | 19.4 | 25.0 | 27.7 | 15.9 | 20.3 | 26.6 | 29.5 |
| PA | 1.298 | 0.045 | 1.388 | 0.053 | 14.4 | 17.9 | 22.9 | 25.3 | 14.8 | 18.8 | 24.5 | 27.3 |
| RI | 1.356 | 0.049 | 1.419 | 0.056 | 13.8 | 17.6 | 23.0 | 25.6 | 15.0 | 19.2 | 25.2 | 28.1 |
| SC | 1.327 | 0.047 | 1.353 | 0.048 | 14.1 | 17.7 | 22.9 | 25.4 | 13.2 | 17.0 | 22.4 | 24.9 |
| SD | 1.243 | 0.039 | 1.216 | 0.038 | 13.1 | 16.2 | 20.6 | 22.7 | 14.7 | 17.6 | 21.9 | 24.0 |
| TN | 1.244 | 0.039 | 1.222 | 0.037 | 13.2 | 16.3 | 20.7 | 22.8 | 12.8 | 15.7 | 19.9 | 21.9 |
| TX | 1.243 | 0.039 | 1.381 | 0.049 | 13.1 | 16.2 | 20.6 | 22.7 | 12.0 | 15.8 | 21.3 | 23.9 |
| UT | 1.347 | 0.049 | 1.339 | 0.048 | 14.3 | 18.1 | 23.5 | 26.0 | 14.0 | 17.7 | 23.0 | 25.5 |
| VT | 1.454 | 0.055 | 1.679 | 0.072 | 12.8 | 17.1 | 23.2 | 26.1 | 13.5 | 19.0 | 26.7 | 30.3 |
| VA | 1.334 | 0.048 | 1.382 | 0.054 | 14.4 | 18.1 | 23.4 | 25.9 | 15.8 | 19.8 | 25.6 | 28.3 |
| WA | 1.243 | 0.039 | 1.307 | 0.045 | 13.1 | 16.2 | 20.6 | 22.7 | 13.6 | 17.1 | 22.1 | 24.5 |
| WV | 1.317 | 0.047 | 1.350 | 0.049 | 14.8 | 18.4 | 23.5 | 26.0 | 14.2 | 18.0 | 23.4 | 25.9 |
| WI | 1.345 | 0.049 | 1.346 | 0.051 | 14.6 | 18.4 | 23.8 | 26.4 | 16.1 | 19.9 | 25.4 | 28.1 |
| WY | 1.243 | 0.039 | 1.356 | 0.050 | 13.1 | 16.2 | 20.6 | 22.7 | 14.2 | 18.1 | 23.5 | 26.1 |

NOTES: This table shows combined federal and state income tax parameters in 2007 with and without the deduction of State and Local Taxes (SALT) from federal taxable income. The table also shows effective tax rates for different levels of Adjusted Gross Income (AGI). Federal taxation includes individual income taxes and the employee portion of payroll (FICA) taxes.

## E.5 Appendix to Section 7.4

Table A.19: State Tax Rates in 1980

| State | $t_n^y$ | $t_n^c$ | $t_n^{corp}$ | $t_n^x$ |
|---|---|---|---|---|
| AL | 2.4 | 4 | 5 | 1.7 |
| AZ | 2.9 | 4 | 10 | 3.3 |
| AR | 2.5 | 3 | 6 | 2 |
| CA | 3.1 | 4.8 | 9.6 | 3.2 |
| CO | 2.2 | 3 | 5 | 1.7 |
| CT | .4 | 7.5 | 10 | 3.3 |
| DE | 4.9 | 0 | 8.7 | 2.9 |
| FL | 0 | 4 | 5 | 2.5 |
| GA | 2.8 | 3 | 6 | 2 |
| HI | 4.8 | 4 | 6.4 | 2.1 |
| ID | 3.2 | 3 | 6.5 | 2.2 |
| IL | 2.2 | 4 | 4 | 1.3 |
| IN | 1.6 | 4 | 3 | 1 |
| IA | 3.1 | 3 | 10 | 10 |
| KS | 2.1 | 3 | 6.8 | 2.3 |
| KY | 3.1 | 5 | 5.8 | 1.9 |
| LA | .8 | 3 | 8 | 2.7 |
| ME | 2.5 | 5 | 6.9 | 2.3 |
| MD | 3.2 | 5 | 7 | 2.3 |
| MA | 4.6 | 5 | 9.5 | 4.7 |
| MI | 3.4 | 4 | 2.3 | .8 |
| MN | 5.7 | 4 | 12 | 4 |
| MS | 1.2 | 5 | 4 | 1.3 |
| MO | 1.9 | 3.1 | 5 | 1.7 |
| MT | 3.2 | 0 | 6.8 | 2.3 |
| NE | 2 | 3 | 4.7 | 1.6 |
| NV | 0 | 3 | 0 | 0 |
| NH | .4 | 0 | 8 | 2.7 |
| NJ | 1.9 | 5 | 7.5 | 2.5 |
| NM | 1.6 | 3.8 | 5 | 1.7 |
| NY | 5.1 | 4 | 10 | 5 |
| NC | 4 | 3 | 6 | 2 |
| ND | 1.5 | 3 | 8.5 | 2.8 |
| OH | 1.5 | 4 | 8 | 2.7 |
| OK | 2.2 | 2 | 4 | 1.3 |
| OR | 4.7 | 0 | 7.5 | 2.5 |
| PA | 2.2 | 6 | 10.5 | 3.5 |
| RI | 2.8 | 6 | 8 | 2.7 |
| SC | 3.5 | 4 | 6 | 2 |
| SD | 0 | 5 | 0 | 0 |
| TN | .4 | 4.5 | 6 | 2 |
| TX | 0 | 4 | 0 | 0 |
| UT | 3 | 4 | 4 | 1.3 |
| VT | 3.1 | 3 | 7.5 | 2.5 |
| VA | 3.1 | 3 | 6 | 2 |
| WA | 0 | 4.5 | 0 | 0 |
| WV | 2.5 | 3 | 6 | 2 |
| WI | 4.4 | 4 | 7.9 | 4 |
| WY | 0 | 3 | 0 | 0 |

NOTES: This table shows state tax rates in 1980 for individual income ($t_n^y$), general sales ($t_n^c$), corporate ($t_n^{corp}$), and sales-apportioned corporate ($t_n^x$) taxes, which is the product of the statutory corporate tax rate and the state's sales apportionment weight.

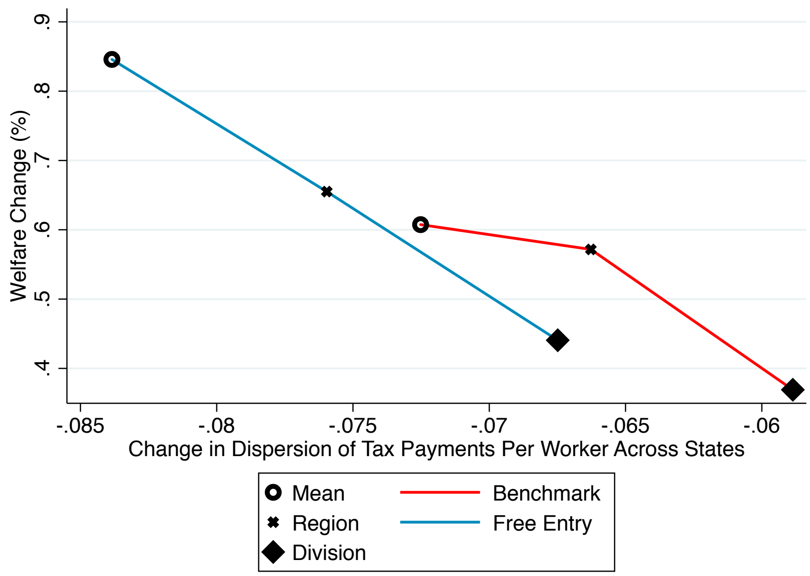
Table A.20: State Income Tax Parameters and Effective Tax Rates in 1980

| State | $a_{n,state}$ | $b_{n,state}$ | State tax rates if AGI is | | | | Overall tax rates if AGI is | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 25K | 50K | 100K | 200K | 25K | 50K | 100K | 200K |
| AL | 1.025 | 0.005 | 2.0 | 2.4 | 3.1 | 3.4 | 14.8 | 18.1 | 22.9 | 25.2 |
| AK | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| AZ | 1.078 | 0.008 | 0.2 | 1.0 | 2.2 | 2.7 | 13.4 | 17.0 | 22.2 | 24.7 |
| AR | 1.092 | 0.011 | 1.2 | 2.1 | 3.6 | 4.3 | 14.1 | 17.9 | 23.3 | 25.9 |
| CA | 1.102 | 0.011 | 0.3 | 1.3 | 2.7 | 3.5 | 13.4 | 17.2 | 22.7 | 25.3 |
| CO | 1.066 | 0.008 | 1.3 | 2.0 | 3.2 | 3.7 | 14.2 | 17.8 | 23.0 | 25.5 |
| CT | 1.087 | 0.010 | 0.9 | 1.8 | 3.2 | 3.9 | 13.9 | 17.7 | 23.0 | 25.6 |
| DE | 1.067 | 0.008 | 1.0 | 1.7 | 2.9 | 3.4 | 14.0 | 17.6 | 22.8 | 25.2 |
| FL | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| GA | 1.132 | 0.014 | 0.8 | 2.1 | 4.0 | 5.0 | 13.8 | 17.9 | 23.7 | 26.4 |
| HI | 1.136 | 0.015 | 1.2 | 2.5 | 4.5 | 5.5 | 14.1 | 18.2 | 24.1 | 26.8 |
| ID | 1.166 | 0.017 | 0.5 | 2.1 | 4.4 | 5.5 | 13.6 | 17.9 | 24.0 | 26.8 |
| IL | 1.019 | 0.004 | 1.4 | 1.8 | 2.3 | 2.5 | 14.3 | 17.6 | 22.3 | 24.6 |
| IN | 1.019 | 0.004 | 2.2 | 2.6 | 3.2 | 3.5 | 15.0 | 18.3 | 23.1 | 25.3 |
| IA | 1.122 | 0.014 | 1.2 | 2.5 | 4.3 | 5.2 | 14.2 | 18.2 | 23.9 | 26.6 |
| KS | 1.066 | 0.009 | 1.6 | 2.4 | 3.6 | 4.2 | 14.5 | 18.1 | 23.3 | 25.8 |
| KY | 1.070 | 0.009 | 1.9 | 2.7 | 4.0 | 4.6 | 14.7 | 18.4 | 23.6 | 26.1 |
| LA | 1.082 | 0.010 | 1.0 | 1.9 | 3.2 | 3.8 | 14.0 | 17.7 | 23.0 | 25.5 |
| ME | 1.131 | 0.015 | 1.2 | 2.5 | 4.5 | 5.5 | 14.1 | 18.2 | 24.0 | 26.8 |
| MD | 1.055 | 0.007 | 1.5 | 2.2 | 3.2 | 3.7 | 14.4 | 17.9 | 23.0 | 25.4 |
| MA | 1.055 | 0.008 | 2.4 | 3.1 | 4.2 | 4.8 | 15.1 | 18.7 | 23.8 | 26.2 |
| MI | 1.049 | 0.007 | 1.5 | 2.1 | 3.1 | 3.5 | 14.4 | 17.9 | 22.9 | 25.3 |
| MN | 1.108 | 0.013 | 1.4 | 2.5 | 4.2 | 5.1 | 14.3 | 18.2 | 23.8 | 26.5 |
| MS | 1.010 | 0.003 | 1.4 | 1.6 | 1.9 | 2.1 | 14.3 | 17.5 | 22.1 | 24.3 |
| MO | 1.065 | 0.008 | 1.3 | 2.0 | 3.1 | 3.7 | 14.2 | 17.8 | 23.0 | 25.4 |
| MT | 1.093 | 0.011 | 1.1 | 2.1 | 3.6 | 4.3 | 14.1 | 17.9 | 23.3 | 25.9 |
| NE | 1.109 | 0.012 | 0.8 | 1.9 | 3.6 | 4.4 | 13.8 | 17.7 | 23.3 | 25.9 |
| NV | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| NH | 1.000 | 0.000 | 0.1 | 0.1 | 0.1 | 0.1 | 13.2 | 16.3 | 20.7 | 22.8 |
| NJ | 1.054 | 0.007 | 0.8 | 1.4 | 2.3 | 2.8 | 13.8 | 17.3 | 22.4 | 24.8 |
| NM | 1.183 | 0.017 | -0.8 | 0.8 | 3.1 | 4.3 | 12.5 | 16.8 | 23.0 | 25.9 |
| NY | 1.099 | 0.012 | 1.3 | 2.4 | 4.0 | 4.7 | 14.3 | 18.1 | 23.6 | 26.2 |
| NC | 1.055 | 0.009 | 2.5 | 3.2 | 4.4 | 5.0 | 15.2 | 18.8 | 23.9 | 26.4 |
| ND | 1.052 | 0.006 | 0.5 | 1.0 | 1.8 | 2.2 | 13.5 | 17.0 | 22.0 | 24.4 |
| OH | 1.061 | 0.008 | 1.2 | 1.9 | 2.9 | 3.4 | 14.1 | 17.7 | 22.8 | 25.2 |
| OK | 1.146 | 0.016 | 0.7 | 2.1 | 4.2 | 5.2 | 13.7 | 17.9 | 23.8 | 26.6 |
| OR | 1.107 | 0.014 | 2.7 | 4.0 | 5.8 | 6.7 | 15.4 | 19.4 | 25.0 | 27.7 |
| PA | 1.046 | 0.007 | 1.5 | 2.1 | 3.0 | 3.5 | 14.4 | 17.9 | 22.9 | 25.3 |
| RI | 1.095 | 0.011 | 0.8 | 1.7 | 3.2 | 3.9 | 13.8 | 17.6 | 23.0 | 25.6 |
| SC | 1.071 | 0.009 | 1.1 | 1.9 | 3.0 | 3.6 | 14.1 | 17.7 | 22.9 | 25.4 |
| SD | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| TN | 1.001 | 0.000 | 0.1 | 0.1 | 0.2 | 0.2 | 13.2 | 16.3 | 20.7 | 22.8 |
| TX | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| UT | 1.087 | 0.011 | 1.4 | 2.3 | 3.8 | 4.5 | 14.3 | 18.1 | 23.5 | 26.0 |
| VT | 1.177 | 0.017 | -0.5 | 1.1 | 3.4 | 4.6 | 12.8 | 17.1 | 23.2 | 26.1 |
| VA | 1.076 | 0.010 | 1.6 | 2.4 | 3.7 | 4.4 | 14.4 | 18.1 | 23.4 | 25.9 |
| WA | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |
| WV | 1.062 | 0.009 | 1.9 | 2.7 | 3.9 | 4.4 | 14.8 | 18.4 | 23.5 | 26.0 |
| WI | 1.086 | 0.011 | 1.8 | 2.8 | 4.2 | 4.9 | 14.6 | 18.4 | 23.8 | 26.4 |
| WY | 1.000 | 0.000 | 0.0 | 0.0 | 0.0 | 0.0 | 13.1 | 16.2 | 20.6 | 22.7 |

NOTES: This table shows state income tax parameters in 1980 as well as effective tax rates for different levels of Adjusted Gross Income (AGI). Tax rates reported in columns 4-7 are state-only, while tax rates in columns 8-11 combine federal and state taxation. Federal taxation includes individual income taxes and the employee portion of payroll (FICA) taxes.

## E.6 Appendix Figures to Section 7.6

Figure A.5: Welfare Change and Dispersion of Tax Payments Under Free Entry




Notes: This figure plots the change in welfare against the change in dispersion of tax liabilities per worker associated with tax harmonization to the national, region, and division means under free entry ($\chi_{FE} = 1$).

# F Data Sources

In this section we describe the data used in sections 3.1, 6, and 7.

## F.1 Government Finances

- State revenue from sales, income and corporate taxes taxes ($R_n^c$, $R_n^y$, $R_n^{corp}$): Source: U.S. Census Bureau – Governments Division; Dataset: Historical State Tax Collections; Variables: corporate, individual, and general sales taxes, which are CorpNetIncomeTaxT41, IndividualIncomeTaxT40, TotalGenSalesTaxT09. We also collect information from the variable TotalTaxes, which include the three types we measure as well as fuels taxes, select sales taxes, and a few other miscellaneous and minor sources of tax revenue.

- State direct expenditures: Source: U.S. Census Bureau – Governments Division; Dataset: State Government Finances; Variable: direct expenditures.

- Effective state individual income tax rate $t_n^y$. Source: SOI Public Use Tax Files (PUFs). Dataset construction: we draw individual taxpayer data from SOI Public Use Tax Files (PUFs) and use NBER's TAXSIM to simulate federal and state individual income tax liabilities under each year's tax law. We drop observations with negative Adjusted Gross Income (AGI) and winsorize observations by their effective state and federal tax rate at the 1st and 99th percentile. We compute $t_n^y$ as the state-year ratio of average AGI over average state income tax liabilities.

- State sales tax rate $t_n^c$: Source: Book of the States; Dataset: Table 7.10 State Excise Tax Rates; Variable: general sales and gross receipts tax (percent).

- State corporate tax rate and apportionment data for $t_n^x$ and $t_n^l$: Source: Suárez Serrato and Zidar (2016).

- Effective federal corporate tax rate $t_{fed}^{corp}$: Source: IRS, Statistics of Income; Dataset: Corporation Income Tax Returns (historical); Variable: effective corporate tax rate = total income tax/ net income (less Deficit); i.e., the effective rate is row 83 divided by row 77.

- Federal individual income tax rate $t_{fed}^y$: Source: NBER TAXSIM; Dataset: Marginal and Average Tax Rates and Elasticities for the U.S., using a fixed 1984 (but in/deflated) sample of taxpayers; Variable: average effective federal tax rate on income, "fed_avg", by state and year.

- Federal payroll tax rate $t_{fed}^w$: Source: Congressional Budget Office; Dataset: Average Federal Tax Rates in 2007; Variable: average payroll tax rates. See Table A.2 for the average in 2007 and additional details in the table notes.

- Income tax schedule parameters $a_n^y$ and $b_n^y$. Source: SOI Public Use Tax Files (PUFs). Dataset construction: we draw individual taxpayer data from SOI Public Use Tax Files (PUFs). In order to abstract from changes in the income distribution over time, we choose 2007 as base year and deflate or inflate nominal variables using the Personal Consumption Expenditure (PCE) index. Then, we use NBER's TAXSIM to simulate state and federal individual income tax liabilities under each year's tax law. We compute individual $t_{fed}^y$ and $t_{n,state}^y$ by dividing federal and state income tax liabilities by Adjusted Gross Income, respectively. We drop observations with negative AGI and winsorize observations by their effective state and federal tax rate at the 1st and 99th percentile. Then, we fit the following models: $1 - t_{fed}^y = a_{fed}^y y^{-b_{fed}^y}$ and $1 - t_{n,state}^y = a_{n,state}^y y^{-b_{n,state}^y}$. In order to estimate $a_{n,state}^y$ and $b_{n,state}^y$, we use OLS to estimate $\gamma$ and $\lambda$ in the regression $\ln(1 - t_{i(n),state,t}^y) = \gamma - \lambda \ln \text{AGI}_{i(n),t} + \varepsilon_{i(n),t}$ for each state-year pair, where $i(n)$ denotes taxpayer $i$ residing in state $n$, and then compute $a_{n,state}^y = e^\gamma$ and $b_{n,state}^y = \lambda$. In order to estimate $a_{fed}^y$ and $b_{fed}^y$, we pool states and follow a similar procedure. Finally, we compute $a_n^y = a_{fed}^y (a_{n,state}^y)^{1-b_{fed}^y}$ and $b_n^y = b_{n,state}^y + b_{fed}^y - b_{n,state}^y b_{fed}^y$.

- Income tax schedule parameters $a_n^y$ and $b_n^y$ when state and local taxes are not deductible from federal taxable income. Source: SOI Public Use Tax Files (PUFs). Dataset construction: we draw individual taxpayer data from SOI Public Use Tax Files (PUFs) and set to 0 three variables associated with state and local tax deductions (*i.e., data50, data51, data52*, which are state and local income taxes, sales taxes, and real estate deductions, respectively). In order to abstract from changes in the income distribution over time, we choose 2007 as base year and deflate or inflate nominal variables using the Personal Consumption Expenditure (PCE) index. Then, we use NBER's TAXSIM to simulate state and federal individual income tax liabilities under each year's tax law. We compute individual $t_{fed}^y$ and $t_{n,state}^y$ by dividing federal and state income tax liabilities by Adjusted Gross Income, respectively. We drop observations with negative AGI and winsorize observations by their effective state and federal tax rate at the 1st and 99th percentile. Because of the non-deductibility of state taxes, we fit the following model: $1 - t_{fed}^y - t_{n,state}^y = a_n^y y^{-b_n^y}$. We use OLS to estimate $\gamma$ and $\lambda$ in the regression $\ln(1 - t_{i(n),fed,t}^y - t_{i(n),state,t}^y) = \gamma - \lambda \ln \text{AGI}_{i(n),t} + \varepsilon_{i(n),t}$ for each state-year pair, where $i(n)$ denotes taxpayer $i$ residing in state $n$, and then compute $a_n^y = e^\gamma$ and $b_n^y = \lambda$.

- Ratio of state and local to state tax revenue for income, sales, and corporate taxes; i.e. $\frac{R_n^{StandLocal,j}}{R_n^{State,j}}$ $\forall j \in \{y, c, corp\}$, respectively. Source: U.S. Census Bureau – Governments Division; Dataset: State and Local Government Finances; Variable: State and Local Revenue; State Revenue (Note that sales taxes uses the general sales tax category)

- We derive the following variables from the sources listed above (for Figure A.1):

  - State and local corporate tax rate: $t_n^{corp,s+l} = t_n^{corp} \times \frac{R_n^{StandLocal,corp}}{R_n^{State,corp}}$.
  - State and local sales tax rate $t_n^{c,s+l} = t_n^c \times \frac{R_n^{StandLocal,c}}{R_n^{State,c}}$.
  - State and local income tax rate $t_n^{y,s+l} = t_n^y \times \frac{R_n^{StandLocal,y}}{R_n^{State,y}}$.

## F.2 Calibration (Section 6.1) and Over-Identification Checks (Section 6.4)

Given the model elasticities and taxes, implementing the counterfactuals in equations (A.24)-(A.46) requires data on $\{s_{in}, \lambda_{in}, L_n, M_n, w_n z_n^L, h_n, P_n Q_n, X_n\}$. We describe here how each of these variables is constructed, alongside other measures used at other states of the quantification.

- Number of Workers $L_n$: Source: 2007 Economic Census of the United States; Dataset: EC0700A1 - All sectors: Geographic Area Series: Economy-Wide Key Statistics: 2007; Variable: Number of paid employees for pay period including March 12

- Annual Hours worked $h_n$: Source: IPUMS; Dataset: March Current Population Survey (CPS); Variable Construction: the number of weeks worked (*wkswork1*) is multiplied by the usual number of hours worked per week (*uhrsworkly*). Sample: our sample is restricted to civilian adults between the ages of 18 and 64. In order to be included in our sample, an individual had to be working at least 35 weeks in the calendar year and with a usual work week of at least 30 hours per week. We also drop individuals who report earning business or farm income.

- State average Hourly Wages $w_{nt}^h(i)$: Source: IPUMS; Dataset: March Current Population Survey (CPS); Variable Construction: annual wage income (*incwage*) is divided by annual hours worked. Sample: the sample for this variable is the same as the one we used to compute state-year average annual hours worked. Top-coded values for years prior to and including 1995 are multiplied by 1.5. These wages are used to construct the variable entering in the quantification $z_n^L w_n$ using equation A.55 and the same steps described in Section D.2.1.

- Total sales $X_n^{Total}$: Source: 2007 Economic Census of the United States; Dataset: EC0700A1 - All sectors: Geographic Area Series: Economy-Wide Key Statistics: 2007; Variable: Employer value of sales, shipments, receipts, revenue, or business done. The source of county-level sales is the U.S. Census Bureau - 2007 Survey of Business Owners and Self-Employed Persons (SBO); Variable: Sales and Receipts, Firms with Paid Employees.

- International Exports $Exports_t^{ROW}$: Source: U.S. Department of Commerce International Trade Administration; Dataset: TradeStats Express - State Export Data; Variable: Exports of NAICS Total All Merchandise to World

- Consumption expenditures $P_nC_n$: Source: U.S. Department of Commerce – Bureau of Economic Analysis (BEA) Regional Data; Dataset: Personal Consumption Expenditures by State; Variable: Personal consumption expenditures

- State GDP $GDP_n$: Source: U.S. Department of Commerce – Bureau of Economic Analysis (BEA) Regional Data; Dataset: GSP NAICS ALL and and GSP SIC ALL; Variable: Gross Domestic Product by State

- Value of Bilateral Trade flow $X_{ni}$: Source: U.S. Census Bureau; Dataset: Commodity Flow Survey; Variable: Value. County-level bilateral trade flows are imputed using weighted shares of state flows, where weights correspond to county shares of state payroll.

- Number of Establishments $M_n$: Source: 2007 Economic Census of the United States; Dataset: EC0700A1 - All sectors: Geographic Area Series: Economy-Wide Key Statistics: 2007; Variable: Number of employer establishments

- We derive the following variables from the primary sources listed above:

  - Value of Intermediate Inputs: $P_nI_n = X_n - GDP_n$
  - Total state spending and revenue: $P_nG_n = R_n = T_n^c + T_n^y + R_n^{corp}$.
  - Total expenditures: $P_nQ_n = P_nG_n + P_nI_n + P_nC_n$
  - Sales from state $n$: $X_n = X_n^{Total} - Exports_n^{ROW}$.
  - Sales to the own state: $X_{ii} = X_i - \sum_n X_{ni}$.
  - Share of sales from $n$ to state $i$: $s_{in} = \frac{X_{in}}{\sum_{i'} X_{i'n}}$.
  - Share of expenditures in $i$ from state $n$: $\lambda_{in} = \frac{X_{in}}{\sum_{n'} X_{in'}}$.

## F.3 Estimation (Section 6.2)

The variables used for estimation are different from those used for the calibration due to data availability. In computing both the calibrated parameters and the counterfactuals, we use the Economic Census measures for wages and employment; the reason being that we collect the sales data from the Economic Census as well. However, the Economic Census is available less frequently than the following data sources, which we use for estimation.

- Number of Workers $L_n$: Source: U.S. Census Bureau; Dataset: County Business Patterns (CBP); Variable: Total Mid-March Employees with Noise; Data cleaning: implemented David Dorn's fixed-point algorithm to impute employment counts in industry-county-year cells that withheld information for confidentiality reasons, and then summed county-level observations by state and year.

- Annual Payroll: Source: U.S. Census Bureau; Dataset: County Business Patterns (CBP); Variable: Annual Payroll

- Number of Establishments $M_n$: Source: U.S. Census Bureau; Dataset: County Business Patterns (CBP); Variable: Total Number of Establishments

- Annual Hours worked $h_n$: Source: IPUMS; Dataset: March Current Population Survey (CPS); Variable Construction: the number of weeks worked (*wkswork1*) is multiplied by the usual number of hours worked per week (*uhrsworkly*). Sample: our sample is restricted to civilian adults between the ages of 18 and 64. In order to be included in our sample, an individual had to be working at least 35 weeks in the calendar year and with a usual work week of at least 30 hours per week. We also drop individuals who report earning business or farm income.

- Wages from CPS $w_n^{CPS}$: Source: IPUMS; Dataset: March Current Population Survey (CPS); Variable Construction: annual wage income (*incwage*) is divided by annual hours worked. Sample: the sample for this variable is the same as the one we used to compute state-year average annual hours worked. Top-coded values for years prior to and including 1995 are multiplied by 1.5.

- Rental prices $r_n$: Source: IPUMS; Dataset: American Community Survey (ACS); Variable: Mean rent; Sample: Adjusted for top coding by multiplying by 1.5 where appropriate

- Price Index $P_n = P_n^{BLS}$; Source: Bureau of Labor Statistics (BLS); Dataset: Consumer Price Index; Variable: Consumer Price Index - All Urban Consumers; Note: Not available for all states. We used population data to allocate city price indexes in cases when a state contained multiple cities with CPI data (e.g., LA and San Francisco for CA's price index)