— Supplementary (Online) Appendix —

Regional Economic Growth Disparities:
A Political Economy Perspective

Tomohito Okabe, t-okabe@ier.hit-u.ac.jp
Timothy Kam, tcy.kam@gmail.com
This document is public access: https://github.com/phantomachine/_regional-growth-politics

Appendix A Proof for Lemma 1

Using Equations (3) and (4), the first-order derivatives of $c_t^{y,k}$ and $c_{t+1}^{o,k}$ with respect to e are given by:

$$\frac{\partial c_t^{y,k}}{\partial e} = -\frac{(1-\tau^\ell)w_t}{1+\beta} < 0,$$
$$\frac{\partial c_{t+1}^{o,k}}{\partial e} = -\frac{\beta(1-\tau^k)(1+r_{t+1})(1-\tau^\ell)w_t}{1+\beta} < 0.$$

Therefore,

$$\frac{\partial U^k}{\partial e} = \frac{1}{c_t^{y,k}}\frac{\partial c_t^{y,k}}{\partial e} + \frac{\partial c_{t+1}^{o,k}}{\partial e} < 0.$$

Thus $U^k(K_t, \cdot)$ is strictly decreasing on $E := [0, 1]$. In addition, given K_t is fixed at the beginning of each date t, $U^\ell(K_t)$ is a strictly positive and constant-valued function. Letting $\mathcal{U}^k$ denote a whole set of $U^k(K_t, \cdot)$, the graphs of $U^k(K_t, \cdot)$ and $U^\ell(K_t)$ obviously have a unique intersection in $(E \times \mathcal{U}^k)$ space. Figure 14 depicts the unique $e^* \in (0, 1)$ in Lemma 1. ■

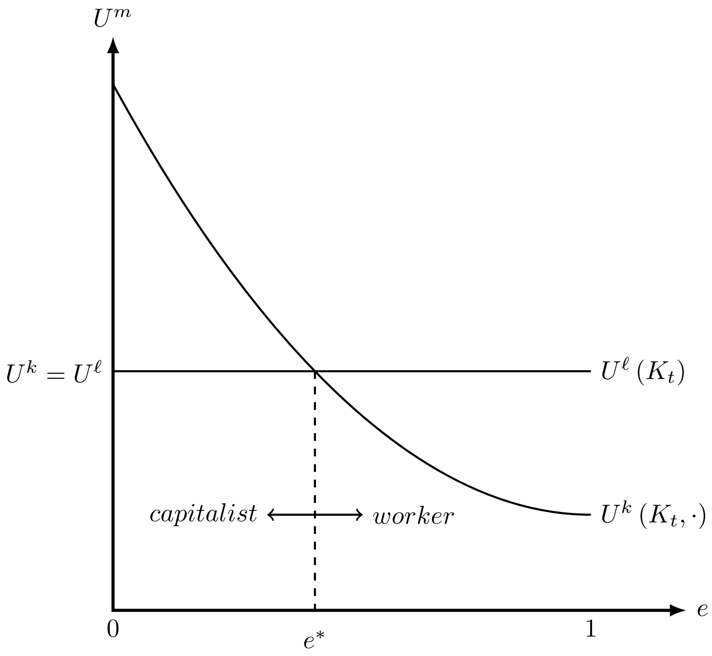


Figure 14: *Ex ante* decision

Appendix B Random e with continuous density

Suppose e is an i.i.d. random variable, and the distribution is given by a smooth density function $h(\cdot)$ on $[0, 1]$. What follows focuses on some derivations that differ from the case of uniform distribution. First, Lemma 1 obviously holds also in this case. The fraction of agents who choose to be capitalists is given by:

$$\mu_t^y = \int_0^{e^*} h(z)dz.$$

The aggregate labour supply by capitalists in this case is given by:

$$\ell_t = N \int_0^{e^*} h(z) dz.$$

Aggregate consumption is:

$$\begin{aligned} C_t &= \int_0^{e^*} c_t^{y,k} h(z) dz + (1-\mu_t^y) c_t^{y,\ell} + \int_0^{e^*} c_t^{o,k} h(z') dz' + (1-\mu_t^o) c_t^{o,\ell} \\ &= (1-\tau^\ell) L_t w_t - (S_t + N\mu_t^y \gamma_t) + (1-\tau^k)(1+r_t) S_{t-1} + N(1-\mu_t^o) c_{g,t}; \end{aligned}$$

Aggregate savings is:

$$S_t = \int_0^{e^*} s_t h(z) dz,$$

where s_t is given by Equation (2). Note that the market clearing condition still holds. In the electoral competition, the expected vote share is now given by:

$$\begin{aligned} \pi_A(K_t, g_t^A, g_t^B) = N \int_0^{\mu_t^o} & h(z) \Psi_t^k \left[V^k \left(K_t, g_t^A, z \right) - V^k \left(K_t, g_t^B, z \right) \right] dz \\ & + N(1-\mu_t^o) \Psi_t^\ell \left[V^\ell \left(K_t, g_t^A \right) - V^\ell \left(K_t, g_t^B \right) \right]. \end{aligned}$$

The first-order condition is then given by:

$$\phi_t^k(0) \int_0^{\mu_t^o} \frac{\partial V^k(z)}{\partial g_t} h(z) dz + (1-\mu_t^o) \phi_t^\ell(0) \frac{\partial V^\ell}{\partial g_t} = 0.$$

Under these characterisations, the political equilibrium still exists. It is straightforward to show that Lemma 2 still holds.

Appendix C Proof for Lemma 2

The first-order derivative of $V^m(\cdot)$ with respect to g_t is given by:

$$\frac{\partial V^m}{\partial g_t} = \beta \frac{1}{c_t^{o,m}} \frac{\partial c_t^{o,m}}{\partial g_t},$$

where

$$\begin{aligned} \frac{\partial c_t^{o,k}}{\partial g_t} &= \frac{\beta \left(1-\tau^k\right) \left[\left(1-\tau^\ell\right) (1-e) w_t - \gamma_t \right]}{1+\beta} \frac{\partial (1+r_t)}{\partial g_t} > 0, \\ \frac{\partial c_t^{o,\ell}}{\partial g_t} &= \frac{1}{N(1-\mu_t^o)} \left[\frac{\partial T_t}{\partial g_t} - 1 \right]. \end{aligned}$$

Therefore, V^k is monotonically increasing, whereas V^ℓ is not. The second order derivatives are given by:

$$\frac{\partial^2 V^m}{\partial g_t^2} = -\beta \frac{1}{(c_t^{o,m})^2} \frac{\partial c_t^{o,m}}{\partial g_t} + \beta \frac{1}{c_t^{o,m}} \frac{\partial^2 c_t^{o,m}}{\partial g_t^2},$$

where

$$
\begin{aligned}
\frac{\partial^2 c_t^{o,k}}{\partial g_t^2} &= \frac{\beta\left(1-\tau^k\right)\left[\left(1-\tau^\ell\right)(1-e)\,w_t-\gamma_t\right]}{1+\beta}\frac{\partial^2\left(1+r_t\right)}{\partial g_t^2}<0,\\
\frac{\partial c_t^{o,\ell}}{\partial g_t} &= \frac{1}{N(1-\mu_t^o)}\left[\frac{\partial^2 T_t}{\partial g_t^2}\right]\\
&= \frac{1}{N(1-\mu_t^o)}\left[\tau^\ell L_t\frac{\partial^2 w_t}{\partial g_t^2}+\tau^k S_t\frac{\partial^2\left(1+r_t\right)}{\partial g_t^2}\right]<0.
\end{aligned}
$$

Therefore,

$$
\frac{\partial^2 V^m}{\partial g_t^2}<0 \quad \forall m\in\{k,l\}.
$$

Given the concavity and the monotonic increasing property, V^k is maximised at a corner point given by:

$$
g_t^{\star k}=T_t.
$$

Likewise, given the concavity, V^ℓ is maximised at a unique point given by

$$
g_t^{\star\ell}\in G_t \quad \text{such that} \quad \left.\frac{\partial V^\ell}{\partial g_t}\right|_{g_t=g_t^{\star\ell}}=0 \qquad \blacksquare
$$

Appendix D Proof for Proposition 1

From Lemma 2, $V^m(\cdot)$, for all $m\in\{k,\ell\}$, is monotonically increasing on $[0,g_t^{\star\ell}]$. Therefore, given K_t, the conflicting preference arises on $(g_t^{\star\ell},T_t)$. The corner point $g_t^{\star\ell}$ is given as the solution of an equation given by:

$$
\begin{aligned}
&\frac{\partial V^\ell}{\partial g_t}=0\\
\Leftrightarrow\quad &\frac{\partial T_t}{\partial g_t}-1=0\\
\Leftrightarrow\quad &\Xi\frac{\partial Y_t}{\partial g_t}-1=0\\
\Leftrightarrow\quad &\eta=\left[\frac{1}{(1-\alpha)A\Xi L_t^{1-\alpha}}\right]^{\frac{1}{\alpha}}
\end{aligned}
$$

Likewise, $g_t^{\star k}=T_t$ is the solution of an equation given by:

$$
\begin{aligned}
&T_t=g_t\\
\Leftrightarrow\quad &\Xi Y_t+\tau^k N\mu_t^y\rho K_t-g_t=0\\
\Leftrightarrow\quad &\Xi A L_t^{1-\alpha}\eta^\alpha+\tau^k\mu_t^y\rho\eta-1=0
\end{aligned}
$$

Summarising these results yields the proposition. $\blacksquare$

Appendix E Estimation of political bias and model calibrations

This section outlines the estimation procedure of the transitional matrix $\mathbf{\Lambda}$ for the Markov chain representing the stochastic process of voters' political bias. Then we discuss how the rest of each model is calibrated to observed long-run data of American states.

E.1 Estimating the distributions of voter political biases ("political factor")

The estimation method follows from Okabe (2015). The computation consists of two steps.

E.2 Multinomial logistic regression

The first step conducts multinomial logistic (MNL) regression for every sample group to compute the sequences of the PID distribution, i.e., $\{\boldsymbol{\phi}\}_{t=0}^{T}$.

Consider the estimation for a sample group. Let a set of independent variables $\mathbf{c}_i$ denote individual i's attributes. $\mathbf{c}_i$ is a $(K+1)$-vector with first-element unity. The utility individual i obtained from choosing alternative ω_j is given by:

$$W_{ij} := \mathbf{c}_i'\boldsymbol{\beta}_j + \varepsilon_{ij},$$

where $\boldsymbol{\beta}_j$ is a $(K+1)$-vector. ε_{ij} is an error term, and follows an i.i.d. type I extreme value distribution. Individual i maximises her utility W_{ij} by choosing alternative d_i. The response probability that individual i chooses PID $\omega_j \in \Omega$ is then given by:

$$\begin{aligned}
\pi_{ij} &:= \text{Prob}\ (d_i = \omega_j | \mathbf{c}_i) && \text{(E.1)}\\
&= \text{Prob}\left(W_{ij} \geq W_{ik}, \quad \text{for all } k \neq j\right)\\
&= \text{Prob}\left(W_{ij} - W_{ik} \leq 0, \quad \text{for all } k \neq j\right)\\
&= \text{Prob}\left(\varepsilon_{ik} - \varepsilon_{ij} \leq \mathbf{c}_i'\boldsymbol{\beta}_j - \mathbf{c}_i'\boldsymbol{\beta}_k, \quad \text{for all } k \neq j\right) && \text{(E.2)}\\
&= \begin{cases} \dfrac{1}{1 + \sum_{r=2}^{3} \exp\left(\mathbf{c}_i'\boldsymbol{\beta}_r\right)}, & j = 1 \\[2ex] \dfrac{\exp\left(\mathbf{c}_i'\boldsymbol{\beta}_j\right)}{1 + \sum_{r=2}^{3} \exp\left(\mathbf{c}_i'\boldsymbol{\beta}_r\right)}, & j = 2, 3 \end{cases} && \text{(E.3)}
\end{aligned}$$

where $0 < \pi_{ij} < 1$, $\sum_j \pi_{ij} = 1$, and ω_1 is the baseline alternative. Equations (E.2) and (E.3) imply that $\varepsilon_{ik} - \varepsilon_{ij}$ follows a logistic distribution.

It is noteworthy that the error terms ε_{ij} represent cognitive fallacy in survey responses. Recall that PID observations are subjective responses based on personal beliefs. For example, it may happen that an Independent respondent (i.e., an individual who actually has no preferable parties) records her PID as Democrat due to lack of awareness.

Define a probability function as:

$$\eta(d_i \mid \mathbf{c}_i; \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_J) = \prod_{j=1}^{3} \left(\pi_{ij}\right)^{\theta_{ij}}, \tag{E.4}$$

where θ_{ij} is a binary indicator given by:

$$\theta_{ij} = \begin{cases} 1 & \text{if individual } i \text{ chooses alternative } \omega_j = d_i \\ 0 & \text{otherwise} \end{cases}$$

The probability function (E.4) maps individual i's probability of observing her actual response from her attributes. Given the sample size of the cross-section data $\mathcal{I}$, the log-likelihood function is then given by:

$$\ln L(\boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_J; \mathbf{d}, \mathbf{c}) = \ln(\eta(d_i)) \tag{E.5}$$

$$= \sum_{i=1}^{\mathcal{I}} \sum_{j=1}^{3} \theta_{ij} log \pi_{ij}, \tag{E.6}$$

where $\mathbf{d} = \{d_1, \cdots, d_{\mathcal{I}}\}'$ and $\mathbf{c} = \{\mathbf{c}_1, \cdots, \mathbf{c}_{\mathcal{I}}\}'$. We can obtain the estimator $\hat{\boldsymbol{\beta}}_j$ by maximising the log-likelihood function (E.6).

After obtaining the estimates for $\boldsymbol{\beta}_j$, we compute the PID distribution for each socio-economic category. Suppose that individual i belongs to category $m \in \mathcal{M}$ (e.g., female, Hispanic, in her 30s). The probability that a m-category individual possesses alternative ω_j is given by

$$\hat{\pi}_{ij} = \frac{\exp\left(\mathbf{c}_i' \hat{\boldsymbol{\beta}}_j\right)}{1 + \sum_{r=2}^{3} \exp\left(\mathbf{c}_i' \hat{\boldsymbol{\beta}}_r\right)} \qquad \text{for } j = 1, 2, 3, \tag{E.7}$$

Using $\hat{\pi}_j$ as estimated proxies for $\phi_t(\omega_j)$ in the Markov chain, we can obtain the sequence of PID distributions $\{\boldsymbol{\phi}_t\}_{t=0}^{T}$.

The values of $\hat{\boldsymbol{\beta}}_j$ would vary across different samples. Nevertheless, it is well known that maximum likelihood estimators asymptotically follow normal distributions; for example, see Wooldridge (2010) for the proof. This allows one to treat the estimates as random variables with asymptotic normal distributions, and hence conduct some significance tests for the estimates, e.g., Wald tests.

The asymptotic normality assumption also allows one to compute confidence intervals. The $100(1-\alpha)\%$ confidence interval for $\hat{\boldsymbol{\beta}}_j$ is given by:

$$\mathbf{c}_i{}' \hat{\boldsymbol{\beta}}_j \pm z_{\alpha/2} \sigma, \tag{E.8}$$

where $z_{\alpha/2}$ is $100(1-\alpha/2)$ percentile point of a standard normal distribution, and σ is the standard error. We can calculate σ from a Hessian matrix in the maximisation problem of the log likelihood (E.6); see, e.g., Wooldridge (2010) for more details. We will utilise this confidence interval in the second step.

E.3 Maximum entropy method

The second step estimates the transitional matrix $\boldsymbol{\Lambda}$ with the maximum entropy method. Entropy is a measure of the statistician's uncertainty regarding a probability distribution parametrised on a set of events. This notion was advocated by Shannon (1948). Applying this measure to estimation, Jaynes (1957a,b) developed the principle of maximum entropy. The key idea is to estimate unknown parameters such that they maximise the entropy subject to imposed constraints on observations and other available information. In practice, we can employ a more general variant of the method proposed by Golan et al. (1996), i.e., the generalised entropy method (GME).

Suppose we obtained the sequence of PID distributions, $\{\boldsymbol{\phi}_t\}_{t=0}^{T}$ in the first step. The observation constraints are given by:

$$\mathbf{b} = \mathbf{A}\mathbf{p} + \mathbf{e},$$

where

$$
\begin{aligned}
&\mathbf{b} := [\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_T]' \\
&\mathbf{A} := \begin{bmatrix} \mathbf{A}(0) \\ \vdots \\ \mathbf{A}(T-1) \end{bmatrix}, \qquad \mathbf{A}(t) := \begin{bmatrix} \boldsymbol{\phi}_t & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\phi}_t & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\phi}_t \end{bmatrix}, \\
&\mathbf{p} := [\boldsymbol{\lambda}(1)', \boldsymbol{\lambda}(2)', \boldsymbol{\lambda}(3)']' = [\lambda_{11}, \lambda_{21}, \lambda_{31}, \lambda_{12}, \lambda_{22}, \lambda_{32}, \lambda_{13}, \lambda_{23}, \lambda_{33}]' \\
&\mathbf{e} := [e_{11}, e_{12}, e_{13}, \cdots, e_{T1}, e_{T2}, e_{T3}]'.
\end{aligned}
$$

where $\mathbf{b}$ is a $3T$-vector of the observations, $\mathbf{A}$ is a $(3T \times 3^2)$ matrix of the observations, $\mathbf{p}$ is the vectorised transitional matrix, $\mathbf{e}$ is a $3T$-vector of unobserved noise. The noise term will be included in the constraint. Recall that the PID distribution $\boldsymbol{\phi}_t$ is computed based on the maximum likelihood estimator $\hat{\boldsymbol{\beta}}_j$ that varies according to a normal distribution. Therefore, $\{\boldsymbol{\phi}_t\}_{t=0}^{T}$ is not a deterministic, but a stochastic path varying with $\hat{\boldsymbol{\beta}}_j$, which we refer to as a "noisy" Markov path.

Next, suppose a compact support for each λ_{xy} that allows for treating it as a random variable. The $\mathbf{p}$ vector is then re-parameterised as:

$$
\mathbf{p} = \mathbf{Z}\mathbf{q} = \begin{bmatrix} \mathbf{Z}(1) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}(2) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}(3) \end{bmatrix} \begin{bmatrix} \mathbf{q}(1) \\ \mathbf{q}(2) \\ \mathbf{q}(3) \end{bmatrix}, \qquad \mathbf{Z}(j) := \begin{bmatrix} \underline{z}_{j1} & \overline{z}_{j1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \underline{z}_{j2} & \overline{z}_{j2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \underline{z}_{j3} & \overline{z}_{j3} \end{bmatrix}
$$

$$
\mathbf{q}(j) := \left[\underline{q}_{j1}, \overline{q}_{j1}, \underline{q}_{j2}, \overline{q}_{j2}, \underline{q}_{j3}, \overline{q}_{j3},\right]'
$$

where $\mathbf{Z}$ is a $(3^2 \times 2 \cdot 3^2)$ sparse matrix representing the supports, $\mathbf{q} \gg \mathbf{0}$ is a $(2 \cdot 3^2)$-vector of probabilities. Likewise, suppose each element of noise vector $\mathbf{e}$ has a compact support. The re-parametrisation is then given by:

$$
\mathbf{e} = \mathbf{V}\mathbf{w} = \begin{bmatrix} \mathbf{V}(0) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}(1) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}(T-1) \end{bmatrix} \begin{bmatrix} \mathbf{w}(0) \\ \mathbf{w}(1) \\ \vdots \\ \mathbf{w}(T-1) \end{bmatrix},
$$

$$
\mathbf{V}(t) := \begin{bmatrix} \underline{v}_{t1} & \overline{v}_{t1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \underline{v}_{t2} & \overline{v}_{t2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \underline{v}_{t3} & \overline{v}_{t3} \end{bmatrix}, \qquad \mathbf{w}(t) := [\underline{w}_{t1}, \overline{w}_{t1}, \underline{w}_{t2}, \overline{w}_{t2}, \underline{w}_{t3}, \overline{w}_{t3}]'
$$

where $\mathbf{V}$ is a $(3T \times 2T \cdot 3)$ sparse matrix representing the supports and $\mathbf{w} \gg 0$ is a $(2T \cdot 3)$ vector of probabilities. The GME estimators are given by solving the following optimisation problem.

$$
\begin{aligned}
&\max_{\mathbf{q},\mathbf{w}} \Gamma(\mathbf{q}, \mathbf{w}) = -\mathbf{q}' \ln(\mathbf{q}) - \mathbf{w}' \ln(\mathbf{w}) \\
&\text{subject to} \\
&\mathbf{b} = \mathbf{A}\mathbf{Z}\mathbf{q} + \mathbf{V}\mathbf{w} \\
&\mathbf{1}_J = (\mathbf{I}_J \otimes \mathbf{1}_2)'\mathbf{q}
\end{aligned}
$$

$$\mathbf{1}_T = (\mathbf{I}_T \otimes \mathbf{1}_2)'\mathbf{w}$$
$$\mathbf{q} \geq \mathbf{0}$$
$$\mathbf{w} \geq \mathbf{0}$$

where Γ is the entropy.

In addition, this paper reflect an extra constraint, $\lambda_{13} = \lambda_{31} = 0$. The transitional matrix to be estimated is then given by:

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & 0 \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ 0 & \lambda_{32} & \lambda_{33} \end{bmatrix}.$$

This follows an hypothesis by Mebane and Wand (1997) that American individual PID gradually changes in the short-run. The hypothesis is verified by a Wald test.

E.4 Results

The data source, the ANES is a repeated cross-section survey. It covers 26 states and the period from 1972 to 2002. We disaggregate the original data into state-by-state samples, and further separate them into eight periods. The data is now divided into (26 states) $\times$ (8 periods) sample groups as shown in Table E.1.

In the first step, we conduct MNL regression, for each sample group, with independent variables for occupations: professional dummy (1 if occupation = professional and managerial, 0 otherwise) and skilled-worker dummy (1 if occupation = skilled, semi-skilled and service workers, 0 otherwise).[19] The baseline PID is Independent. Based on the MNL estimators, we compute the sequence of PID distributions for professional and skilled workers. Table E.2 shows the estimated distributions for Alabama. Some of the distributions are identical due to failures in rejecting the null hypothesis $H_0 := \hat{\boldsymbol{\beta}}_j = \mathbf{0}$ for $j = 2, 3$ at the Wald test. The significance level for the test is 10%.

Given the time sequence of the PID distributions, we estimate the transitional matrix. Table E.3 summarises the estimated transitional matrices, stationary distributions and other relevant results. The Wald test with the null hypothesis $H_0 := \lambda_{13} = \lambda_{31} = 0$ results in failing to reject the null at less than 1% significance level. The resulting stationary distribution give calibration values for the political bias parameters: the fraction of professionals with Independent status, i.e., 0.290 for $\phi^{*k}(0)$, the fraction of skilled workers with Independent status, i.e., 0.319 for $\phi^{*\ell}(0)$.

We conduct the estimation for the other states in a similar fashion, but the entire cross-state results are omitted for brevity.

E.5 Calibration of other model parameters and factors

Period configuration. The model period is set to four years, so that it equals the frequency of the PID data. The active duration of agents is set to eight model periods (i.e., 32 years) that covers the total time period of the PID data (i.e., 30 years). The duration of youth and old age is set to four model periods (i.e., 16 years) each. Thus, this period set-up ensures that the lifetime bias transitions of the model are approximated by the data spanning almost the same range.

In addition, we suppose that individuals are "born" (i.e., the enter the scene) at age 18. They become "old" in four periods (i.e., at age 34), and become inactive after further four periods (i.e., at age 46). The relevant lifetime

[19] We do not employ other independent variables to avoid the multicollinearity that is likely to occur with the following reasons: The sample size of each American state is not necessarily large. Those with the occupation category of professional and managerial have limited population share in each state.

Table E.1: ANES cross-section data

State	1972/ 1974	1976/ 1978	1980/ 1982	1984/ 1986	1988/ 1990	1992/ 1994	1996/ 1998	2000/ 2002	Total
AL	49	86	53	93	113	74	110	100	**678**
AR	170	115	78	92	97	67	23	56	**698**
CA	331	390	278	512	416	412	215	329	**2,883**
CO	48	69	51	76	81	76	62	72	**535**
CT	85	66	54	82	73	49	29	35	**473**
FL	120	184	147	176	117	194	146	150	**1,234**
GA	142	104	79	192	184	222	113	62	**1,098**
IL	208	206	136	119	136	131	75	108	**1,119**
IN	60	102	65	114	82	169	141	74	**807**
IA	103	88	50	87	83	53	51	55	**570**
MD	88	93	42	81	84	87	66	70	**611**
MA	148	135	79	87	76	108	53	101	**787**
MI	177	199	141	299	204	250	126	130	**1,526**
MN	92	85	46	114	101	143	97	115	**793**
MS	121	103	77	72	54	76	52	40	**595**
NJ	150	147	101	117	87	146	84	76	**908**
NY	252	266	197	334	312	268	175	206	**2,010**
NC	203	153	101	120	125	74	24	52	**852**
OH	284	254	182	163	183	118	80	134	**1,398**
OR	73	88	51	93	95	74	70	62	**606**
PA	246	271	157	99	122	116	98	105	**1,214**
TN	53	80	54	187	232	137	71	71	**885**
TX	119	201	166	300	283	329	279	240	**1,917**
VA	75	97	77	147	97	229	227	158	**1,107**
WA	85	101	63	65	88	72	68	85	**627**
WI	37	68	38	97	125	79	134	130	**708**
Total	**3,519**	**3,751**	**2,563**	**3,918**	**3,650**	**3,753**	**2,669**	**2,816**	**26,639**

Table E.2: Computed PID distributions for Alabama

Professionals				
		Democrat	Independent	Republican
Period 1	*	0.531	0.204	0.265
Period 2	*	0.651	0.233	0.116
Period 3		0.769	0.000	0.231
Period 4	*	0.613	0.075	0.312
Period 5		0.417	0.000	0.583
Period 6	*	0.595	0.081	0.324
Period 7		0.318	0.136	0.545
Period 8		0.389	0.000	0.611
Skilled workers				
		Democrat	Independent	Republican
Period 1	*	0.531	0.204	0.265
Period 2	*	0.651	0.233	0.116
Period 3		0.800	0.100	0.100
Period 4	*	0.613	0.075	0.312
Period 5		0.476	0.071	0.452
Period 6	*	0.595	0.081	0.324
Period 7		0.680	0.120	0.200
Period 8		0.731	0.115	0.154

Note that * indicates that the distribution is identical between two groups. This is the case where the Wald test for the whole regressors fails to reject the null at 10% singificance level

Table E.3: Estimated transitional matrices for Alabama

Professionals			
PID_{t-1}	PID_t		
	Democrat	Independent	Republican
Democrat	0.730	0.270	0.000
Independent	0.366	0.276	0.359
Republican	0.000	0.328	0.672
Stationary distribution	0.393	0.290	0.317
$\overline{\Gamma}(\mathbf{q})=$	0.694		
$v=$	[-0.769	0.769]	
Skilled workers			
PID_{t-1}	PID_t		
	Democrat	Independent	Republican
Democrat	0.722	0.278	0.000
Independent	0.376	0.288	0.366
Republican	0.000	0.328	0.672
Stationary distribution	0.432	0.319	0.250
$\overline{\Gamma}(\mathbf{q})=$	0.713		
$v=$	[-0.900	0.900]	

where $\overline{\Gamma}(\mathbf{q})$ is the normalised entropy, v the support of the noise.

of agents is therefore scaled to 46 years, which is not the same as actual life expectancy, but this suffices for the purposes of our accounting exercises. The frequency setting results in eight cohorts at every period. Figure E.1 depicts the generation structure.

Other parameters. Table E.4 shows the calibrated values of parameters we take to be identical across regions. These refer to parameters defining the production and capital-accumulation technologies, and also preferences.

The capital share of output, α, and the depreciation rate, δ, are taken from Cooley and Prescott (1995), but the latter is adjusted to the four-year basis. The TFP parameter, A, is set such that the capital-public good ratio η^* takes reasonable values satisfying constraint (19). The discount factor, β, is simply taken from an empirical study

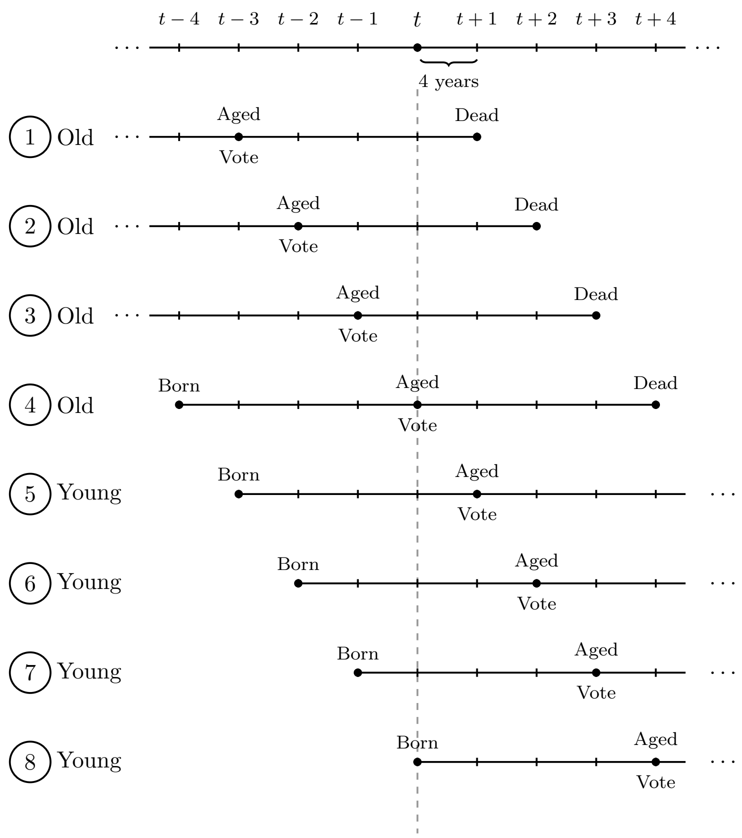


Figure E.1: Eight cohorts at time t

Table E.4: Calibrated region-common parameters

Parameter	Definition	Value	Reference/ process
α	capital share	0.36	Cooley and Prescott (1995)
$1-(1-\delta)^{1/4}$	depreciation rate	0.048	Cooley and Prescott (1995)
A	TFP	3.11	set to match with reference values of capital-output ratio
$\beta^{1/16}$	discount factor	1.011	Hurd (1989)

by Hurd (1989). This is justified by the life-cycle model literature, e.g., Heer and Maussner (2009, Ch7).[20]

Table E.5 shows the calibrated values for tax rates and the steady-state rental rate of capital. The cross-state tax rates have five variations, whereas the rental rates differ across the states. Since no studies estimate state-level effective taxes in the U.S., this paper utilises the national level estimates by Gomme and Rupert (2007) as reference values. The reference can be thought of as the average of the cross-state rates. Group A takes the reference values, and the other four groups (Group B to Group E) take values that are slightly adjusted from the reference. The adjustment for the latter four groups will be conducted to maintain consistency with endogenous outcomes and other parameters. We calibrate the rental rate of capital using Equation (11). However, we will compute the capital-output ratio with reference values, because once again there are no studies that estimate state-level rental rates. The cross-state capital stocks are estimated by Yamarik (2013), but we do not directly employ his estimates because they result in very high rental rates. Instead, we compute the cross-state fractions of the state-level to the national-level estimate, and multiply them by national-level fixed capital assets measured by *the Bureau of Economic*

[20]The literature says that we should empirically measure the discount factor for the OLG model because we have no theoretical restrictions to pin down its value from other observables, unlike in the case of infinite time horizon models.

Analysis (BEA). In turn, what are obtained are the reference values for cross-state capital. The outputs are simply taken from measures of Gross State Product (GSP) available from the BEA. Given the computed capital-output ratios for the period from 1972 to 2003, we obtain the rental rates of capital from 6.0 to 9.1% per annum which are close to the national-level (pre-tax) real return on capital in the business cycle literature; see Gomme and Rupert (2007). Table E.6 shows the calibrated region-specific parameters.

Table E.5: Calibrated tax rates and steady-state rental rate of capital

State		τ^k	τ^ℓ	r^* (annual basis)
Group A				
Arkansas	AR	0.2921	0.2241	0.069
California	CA	0.2921	0.2241	0.071
Connecticut	CT	0.2921	0.2241	0.072
Georgia	GA	0.2921	0.2241	0.072
Illinois	IL	0.2921	0.2241	0.066
Indiana	IN	0.2921	0.2241	0.070
Massachusetts	MA	0.2921	0.2241	0.074
Michigan	MI	0.2921	0.2241	0.073
Minnesota	MN	0.2921	0.2241	0.068
New Jersey	NJ	0.2921	0.2241	0.072
Ohio	OH	0.2921	0.2241	0.071
Oregon	OR	0.2921	0.2241	0.068
Wisconsin	WI	0.2921	0.2241	0.074
Group B				
Colorado	CO	0.2500	0.2000	0.064
New York	NY	0.2500	0.2000	0.063
Pennsylvania	PA	0.2500	0.2000	0.062
Texas	TX	0.2500	0.2000	0.060
Group C				
Alabama	AL	0.2921	0.2400	0.075
Florida	FL	0.2921	0.2400	0.079
Iowa	IA	0.2921	0.2400	0.075
Mississippi	MS	0.2921	0.2400	0.076
Tennessee	TN	0.2921	0.2400	0.078
Virginia	VA	0.2921	0.2400	0.087
Group D				
North Carolina	NC	0.2921	0.2600	0.082
Washington	WA	0.2921	0.2600	0.079
Group E				
Maryland	MD	0.2921	0.2800	0.091

Appendix F Marginal effects of the three growth channels

Previously we showed that the balanced-growth equilibrium path of each regional economy is characterised by a system of four non-linear equations: (22), (23), (24) and (25). From this system of equations, we will be able to extract (implicitly) the growth accounting insights. Before we arrive at this growth accounting exercise proper, we will need to know the marginal contributions of the three channels in each regional model—capital (ρ), labor (θ) and politics ($\phi^{*k}(0)$)—to the four endogenous variables ν, η^*, μ^* and L^*, denoting growth rate, capital-to-public-infrastructure ratio, measure of capitalists, and total labor, respectively. As an illustration, we present the numerical results for the state of Alabama. The same conclusions hold true for all the other states.

Figure F.1 shows the results for Alabama. It consists of nine panels: (A1)-(A4) show the effects of ρ; (B1)-(B4) the effects of $\phi^{k*}(0)$, holding $\phi^{\ell*}(0)$ constant; and (C1)-(C4) the effects of θ. The observed effects are all monotonic with respect to the parameter, except for Panel (C3).

Table E.6: Calibrated region-specific parameters

State		Pecuniary cost parameter ρ	Fraction of neutral capitalists $\phi^{*k}(0)$	Fraction of neutral workers $\phi^{*\ell}(0)$	Time efficiency parameter θ	Size of population N
Group A						
Arkansas	AR	0.077	0.307	0.344	0.162	2.101
California	CA	0.134	0.147	0.175	0.065	2.284
Connecticut	CT	0.019	0.365	0.349	0.281	1.881
Georgia	GA	0.105	0.242	0.281	0.090	2.206
Illinois	IL	0.017	0.262	0.295	0.315	1.847
Indiana	IN	0.103	0.268	0.292	0.133	2.180
Massachusetts	MA	0.117	0.218	0.295	0.042	2.225
Michigan	MI	0.135	0.306	0.263	0.135	2.223
Minnesota	MN	0.027	0.269	0.306	0.249	1.934
New Jersey	NJ	0.078	0.284	0.294	0.159	2.095
Ohio	OH	0.116	0.298	0.301	0.127	2.205
Oregon	OR	0.010	0.183	0.183	0.357	1.787
Wisconsin	WI	0.137	0.260	0.301	0.057	2.278
Group B						
Colorado	CO	0.090	0.301	0.296	0.045	2.148
New York	NY	0.033	0.265	0.150	0.325	1.814
Pennsylvania	PA	0.069	0.274	0.187	0.176	2.022
Texas	TX	0.048	0.254	0.200	0.183	1.995
Group C						
Alabama	AL	0.102	0.290	0.319	0.145	2.166
Florida	FL	0.148	0.182	0.182	0.093	2.293
Iowa	IA	0.105	0.284	0.316	0.139	2.178
Mississippi	MS	0.099	0.213	0.213	0.174	2.122
Tennessee	TN	0.126	0.293	0.318	0.102	2.241
Virginia	VA	0.155	0.277	0.215	0.098	2.251
Group D						
North Carolina	NC	0.069	0.312	0.318	0.243	1.992
Washington	WA	0.078	0.255	0.305	0.202	2.071
Group E						
Maryland	MD	0.121	0.283	0.284	0.174	2.158

Private capital market channel. From Panels (A1)-(A4), the effects of marginal change in ρ are summarised as follows:

$$\frac{\partial \eta^*}{\partial \rho} > 0, \quad \frac{\partial \mu^*}{\partial \rho} < 0, \quad \frac{\partial L^*(\mu^*(\rho))}{\partial \rho} < 0, \quad \frac{\partial \nu\left(\eta^*(\rho), \mu^*(\rho), L^*(\rho), \rho\right)}{\partial \rho} < 0.$$

We derive the economic intuitions behind these effects.

First, recall that $\rho \times K_t$ is the total per-period fixed cost for an agent being a capitalist. Since K_t is fixed at the beginning of each date, an increment in ρ decreases saving, thus having the direct impact of decreasing growth. On the left side of Equation (22), this shows up as an explicit negative ρ term.

Second, observe that the growth rate also depends on three endogenous variables, $\eta^*(\rho), \mu^*(\rho)$ and $L^*(\mu^*(\rho))$. For now, we restrict our attention on the marginal change in the capital-public-good ratio η^* and the agent distribution μ^*. These two variables are determined by Equations (23) and (24). The former is the first-order condition describing a probabilistic voting equilibrium, whereas the latter is the cut-off condition for an agent's *ex-ante* choice. Suppose ρ is marginally increased. In the *ex-ante* choice, on one hand, an increment in ρ reduces the net amount of savings, and hence the *ex-post* payoff of "capitalists", holding the electoral outcome constant. On the other hand, it decreases the capital tax base, and hence the lump sum transfer, thus decreasing the *ex-post* payoff of "workers". Since the lifetime payoffs of both agents are lowered, it is uncertain whether the proportion of capitalists μ^* is eventually increased or decreased. However, this has consequences on the electoral competition that subsequently occurs. In the election, the marginal change in μ^* affects the electoral outcome, summarised

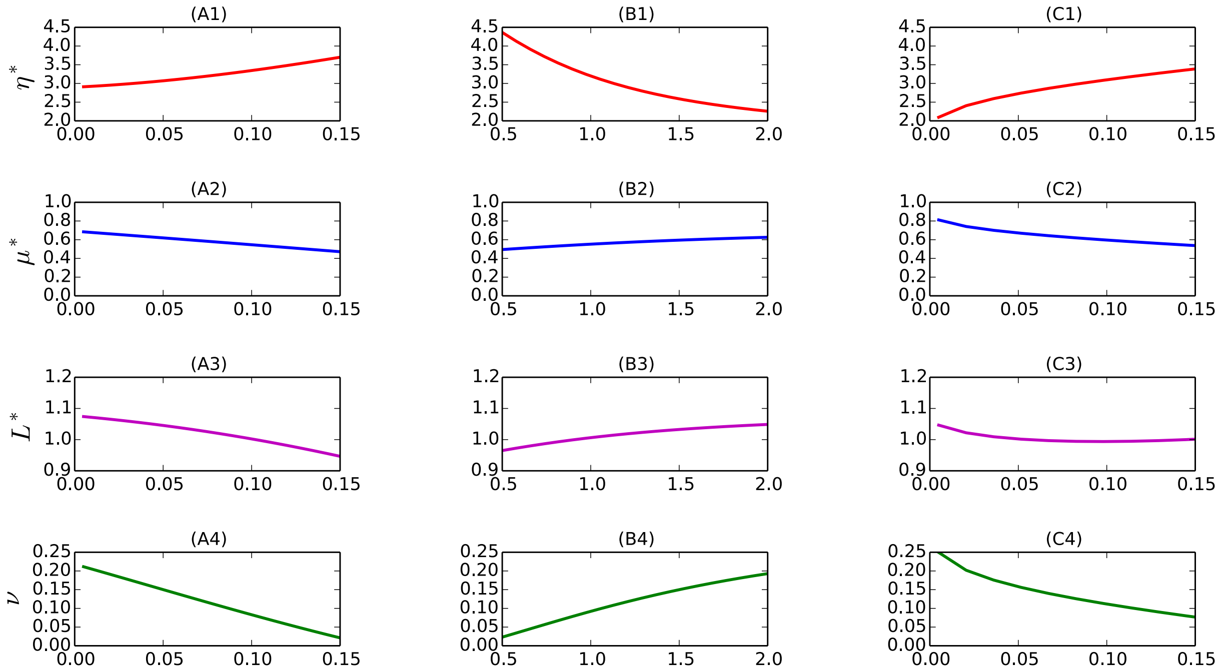


Figure F.1: Effects of parameters for Alabama

in terms of capital-to-public-infrastructure ratio, η^*. Recalling that μ^* is the weight for the welfare of capitalists, an increase in μ^* lowers η^*, whereas a decrease in μ^* raises η^*. In fact, the effect of the electoral outcome feeds back to the *ex-ante* choice because individuals have perfect foresight on the aggregate-level equilibrium outcomes. The numerical comparative statics shows that these endogenous effects eventually result in an increase in η^* and a decrease in μ^*. An increase in η^* is equivalent to a decrease in public good expenditure. A decrease in μ^* lowers savings, and hence decrease capital stock. Both these effects reduce the growth rate.

Finally, we discuss how L^* responds at the margin to a small change in in ρ. From Equation (25), an increase in μ^* raises L^* for small enough θ. This indicates that the labour productivity of capitalists is much larger than that of workers. An increase in the population of capitalists increases the aggregate labour, *viz.* that $L^*(\cdot)$ is an increasing function of μ^*. Therefore, the direction of the marginal effect of ρ on L^* is identical to that on μ^*. A decrease in L^* raises the marginal product of labour, and hence savings. This raises the growth rate, but this positive effect is cancelled out by the negative effect mentioned earlier.

Given these direct and indirect effects, the growth rate monotonically decreases as ρ increases.

Political factor. From Panels (B1)-(B4) in Figure F.1, the effects of the political factor are summarised as:

$$\frac{\partial \eta^*}{\partial \phi^{*k}(0)} < 0, \quad \frac{\partial \mu^*}{\partial \phi^{*k}(0)} > 0, \quad \frac{\partial L^*(\mu^*(\phi^{*k}(0)))}{\partial \phi^{*k}(0)} > 0,$$
$$\frac{\partial \nu\left(\eta^*(\phi^{*k}(0)), \mu^*(\phi^{*k}(0)), L^*(\phi^{*k}(0))\right)}{\partial \phi^{*k}(0)} > 0.$$

We now explain the forces underlying these partial derivative results. Suppose that $\phi^{*k}(0)$ is marginally increased. Since the political bias parameter is not present in the function $\nu(\cdot)$, the change in $\phi^{*k}(0)$ does not directly affect growth, but $\phi^{*k}(0)$ directly affects electoral outcome. In the probabilistic voting equilibrium, $\phi^{*k}(0)$ serves as a weight for the welfare of capitalists. Therefore, an increase in $\phi^{*k}(0)$ lowers η^*, and hence raises the public good expenditure. This increases wages and rent of capital. Those effects are taken into account in agents' *ex-ante* choice. In the *ex ante* choice, the higher factor rental rates raises the *ex-post* payoff of capitalists and workers, thus altering μ^*. The change in μ^* again influences the outcome of the subsequent electoral competition. Similar to that in the capital channel discussed earlier, the numerical results indicate that this makes capitalists much better off than workers, thus increasing μ^*. An increase in μ^* raises the capital stock. An increase in $\phi^{*k}(0)$ enhances the public good expenditure. These two effects increase growth. L^* rises due to increases in μ^*, thus lowering the marginal product of labour. However, this negative effect on growth is dominated by the two positive forces above. In sum, the growth rate is increased by increasing $\phi^{*k}(0)$.

Labour channel. From Panels (C1)-(C4) in Figure F.1, the marginal effects of the labor productivity parameter θ are given by:

$$\frac{\partial \eta^*}{\partial \theta} > 0, \quad \frac{\partial \mu^*}{\partial \theta} < 0, \quad \frac{\partial L^*}{\partial \theta} \gtreqless 0, \quad \frac{\partial \nu\left(\eta^*(\theta), \mu^*(\theta), L^*(\theta)\right)}{\partial \theta} < 0.$$

We explain the forces behind these partial derivative results. Suppose that θ is marginally increased. Similar to the situation for the political factor above, θ does not directly affect growth. We therefore focus on the changes in the endogenous variables. In the *ex-ante* choice, on one hand, an increment in θ raises the labour earnings of workers, thus increasing the *ex-post* payoff of workers. On the other hand, an increase in θ raises the labour tax base. This enhances the productive public good allocation, thus increasing the *ex ante* payoff of capitalists, holding the electoral outcome constant. This changes μ^*, but the effect is ambiguous. The change in μ^* alters the electoral outcome η^*. However, agents will anticipate this electoral outcome change, in their *ex-ante* choice problem. The calibrated and computed results suggest that the overall, we have an increase in η^* and a decrease in μ^*. The change in L^* is not monotonic because the change in θ directly affects the aggregate labour allocation. Given these effects, the growth rate monotonically decreases as the labor productivity parameter increases. This implies that the non-monotonic effect via a change in L^* is not dominant overall.

Appendix G Robustness of channels

Since the ANES data employed in this paper are from a national-level survey, the sample size of the cross-section is not necessarily large. We therefore combined the ANES data with another national-level public opinion survey, the *General Social Survey* (GSS). Combining the data increases the total number of observations from 26,639 to 64,378. We then re-estimated the bias parameters. Table G.1 shows the results in terms of the relative ratio of bias parameters, $\phi^{*k}(0)/\phi^{*\ell}(0)$ The differences in the relative ratios between two estimations vary $0 - 74\%$.

We re-calibrate the models based on these re-estimated parameters, and re-do the growth accounting exercise. Table G.2 shows the results. The region-common parameters are held the same as in Section 3.3. The first six columns summarise the cross-region averages and the standard deviations of the sensitivity measurements. In spite of an alternative data source and the resulting re-calibration, the results do not differ greatly from what we did in the main paper. The most influential channels are: the capital channel (in 20 states), the political (or voter bias distributions) factor (in four states) and the labour channel (in two states).

Table G.1: Estimated bias parameters

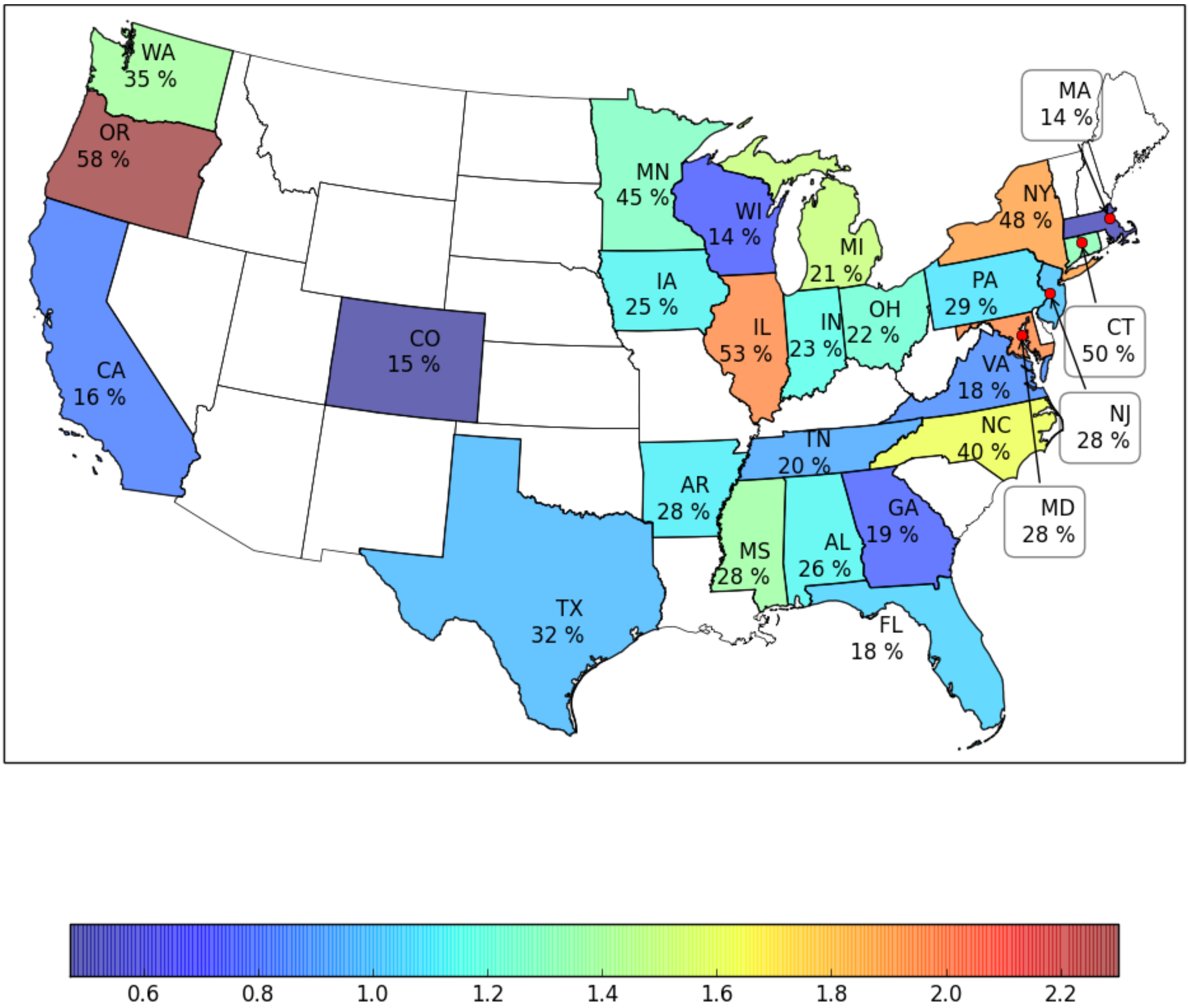
State		$\phi^{*k}(0)/\phi^{*\ell}(0)$ ANES	ANES + GSS	Difference (%)
Arkansas	AR	0.892	0.858	3.9
California	CA	0.840	0.719	14.4
Connecticut	CT	1.046	1.029	1.6
Georgia	GA	0.861	0.963	-11.8
Illinois	IL	0.888	1.232	-38.7
Indiana	IN	0.918	1.406	-53.2
Massachusetts	MA	0.739	1.000	-35.3
Michigan	MI	1.163	0.947	18.6
Minnesota	MN	0.879	1.414	-60.9
New Jersey	NJ	0.966	0.882	8.7
Ohio	OH	0.990	1.087	-9.8
Oregon	OR	1.000	1.000	0.0
Wisconsin	WI	0.864	1.375	-59.2
Colorado	CO	1.017	0.978	3.8
New York	NY	1.767	0.867	50.9
Pennsylvania	PA	1.465	1.358	7.3
Texas	TX	1.270	0.825	35.1
Alabama	AL	0.909	1.584	-74.2
Florida	FL	1.000	1.085	-8.5
Iowa	IA	0.899	0.760	15.4
Mississippi	MS	1.000	0.984	1.6
Tennessee	TN	0.921	1.000	-8.5
Virginia	VA	1.288	0.925	28.2
North Carolina	NC	0.981	1.655	-68.7
Washington	WA	0.836	1.000	-19.6
Maryland	MD	0.996	1.000	-0.4

Table G.2: The results of growth accounting

		ANES	ANES + GSS
R_ρ	Avg	34.2	42.7
	Std. Dev	13.7	12.7
$R_{\phi^{*k}(0)}$	Avg	36.8	37.0
	Std. Dev	2.8	2.8
R_θ	Avg	29.0	20.3
	Std. Dev	12.7	12.3
Most influential channel (Number of states)			
ρ		15	20
$\phi^{*k}(0)$		4	4
θ		7	2

Appendix H Effect of "labour channel" on growth

Figure H.1 shows the effect of θ on growth across states. The colour codes indicate that the effect on growth is large especially in Oregon.



Notes:
(1) Colour codes indicate the level of $\frac{\theta}{\nu}\frac{d\nu}{d\theta}$.
(2) Percentages below state names are ratio R_θ.

Figure H.1: Impact of the labour channel (indexed by θ) on growth

Appendix I Benchmark social planner's problem

The social planner problem is given by:

$$\max_{g_t} \int_o^{\mu_t^y} V^{y,k}(g_t;z)dz + (1-\mu_t^y)V^{y,\ell}(g_t) + \int_0^{\mu_t^o} V^{o,k}(g_t;z')dz' + (1-\mu_t^o)V^{o,\ell}(g_t),$$

where $V^{i,m}(\cdot)$ is the indirect utility function of i-aged, m-type individuals over the policy g_t. Since the objective function is concave, we can obtain a unique solution g_t^*. Notice that the bliss point of young agents is identical to that of old capitalists given by Lemma 2 because an extra increment in g_t raises the current wage, thus raising the earnings of the young. Replacing the political constraint (20) with the first-order condition of the social planner problem allows for obtaining the equilibrium of the benchmark economy. The first-order condition is given by:

$$\int_o^{\mu^*} \frac{\partial V^{y,k}(z)}{\partial g_t}dz + (1-\mu^*)\frac{\partial V^{y,\ell}}{\partial g_t} + \int_0^{\mu^*} \frac{\partial V^{o,k}(z')}{\partial g_t}dz' + (1-\mu^*)\frac{\partial V^{o,\ell}}{\partial g_t} = 0. \tag{I.1}$$

Thus, the steady state of the social planner economy is characterised by a system of Equations (22), (24), (25) and (I.1).

Appendix J Additional results for Section 4

Figure J.1 and J.2 show the political distortion on the welfare of young agents across states. The welfare loss of young capitalists varies by $9-36\%$. The welfare loss of young workers varies by $8-31\%$.

Figures J.3 and J.4 show the distortion in the welfare of young agents due to political bias across states. The welfare benefit of young capitalists varies by -11 – 20%. The welfare benefit of young workers varies by -10 – 19%. Notice the negative values indicating the loss of welfare.

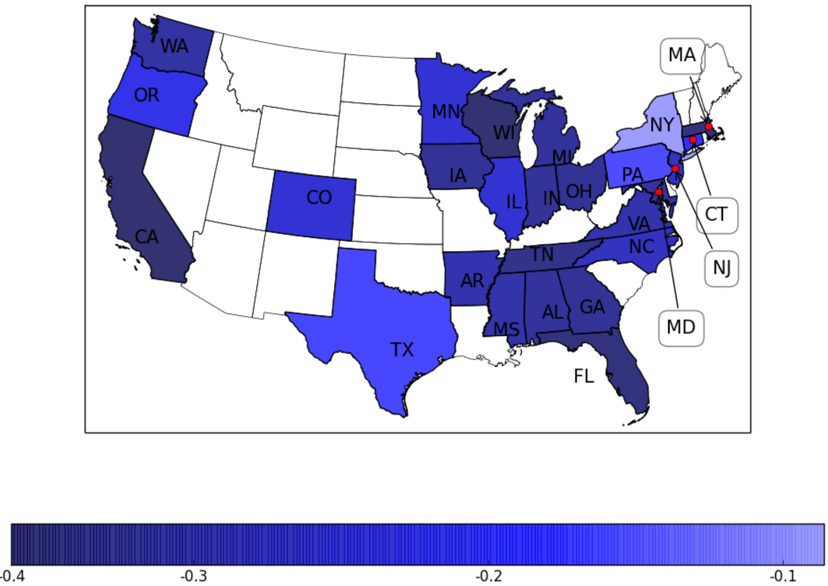
The cross-state patterns for the political distortion and the political bias distortion are the same as Figure 12 and Figure K.2, respectively.

Appendix K Distortion due to political influence

In Section 4, we examined the thought experiment: What is to be lost if the regions moved from an efficient planning outcome to their current politico-economic outcome? Here, we follow up on this theme, and consider the question: What is to be lost if the regions moved from the same political regime (but with equal political influence between capitalists and workers) to its current incarnation where there is asymmetric political influence between capitalist- and worker-voters?

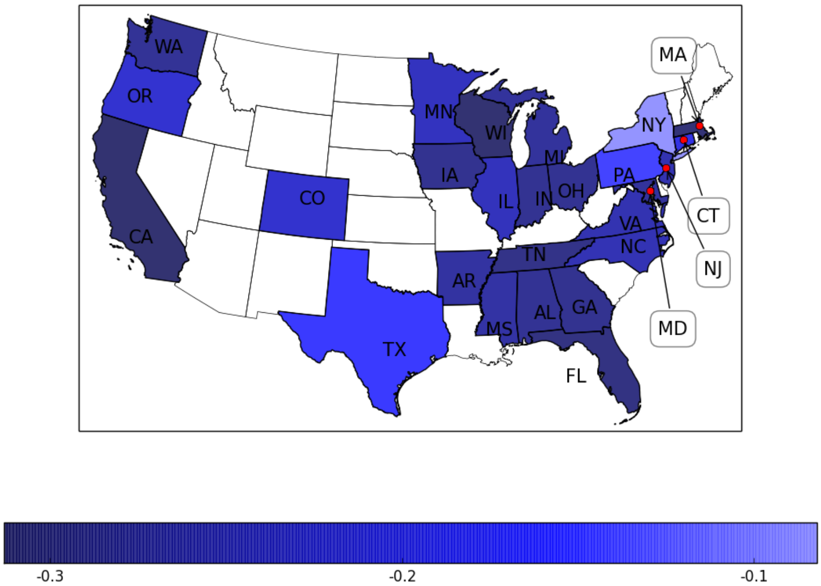
The benchmark is now the political economy with $\phi^{*k}(0)/\phi^{*\ell}(0) = 1$, holding all else identical to the calibrated setting in the paper.

We start with the loss measured in terms of economic growth, $\Delta\nu_{NB} := \nu - \nu_{NB}$, where ν_{NB} denotes the growth rate of the benchmark economy. The results are shown in Figure K.1. No distortion emerges in Florida, Mississippi and Oregon, where the original political economy also possesses no asymmetry in political influence; i.e., $\phi^{*k}(0)/\phi^{*\ell}(0) = 1$, and hence is identical to the benchmark economy. In six states, Connecticut, Michigan, Virginia, Colorado, New York, Pennsylvania and Texas, for which the original parameter $\phi^{*k}(0)/\phi^{*\ell}(0) > 1$, the bias towards capitalists raises the infrastructure expenditure, thus leading to a positive impact on growth. Recall that $\phi^{*k}(0)/\phi^{*\ell}(0)$ serves as the relative weight for the policy favoured by old capitalists in the probabilistic voting, i.e., the infrastructure good allocation. Therefore, a political economy with the increased relative weight enhances infrastructure expenditure. In the rest of the states with the original parameter $\phi^{*k}(0)/\phi^{*\ell}(0) < 1$, the bias occurs



Note that colour codes indicate $\Delta\tilde{c}^{y,k}_{SP}$.

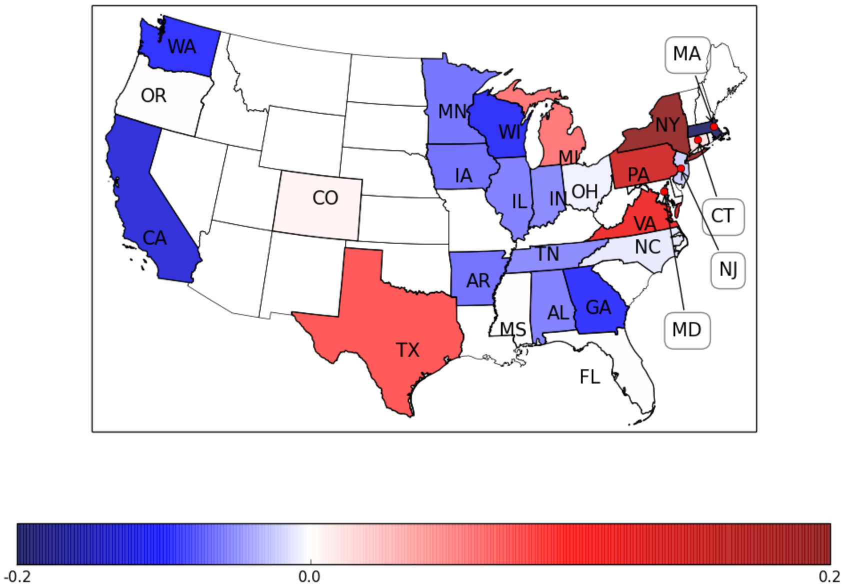
Figure J.1: Political distortion in young capitalists' welfare



Note that colour codes indicate $\Delta\tilde{c}^{y,\ell}_{SP}$.

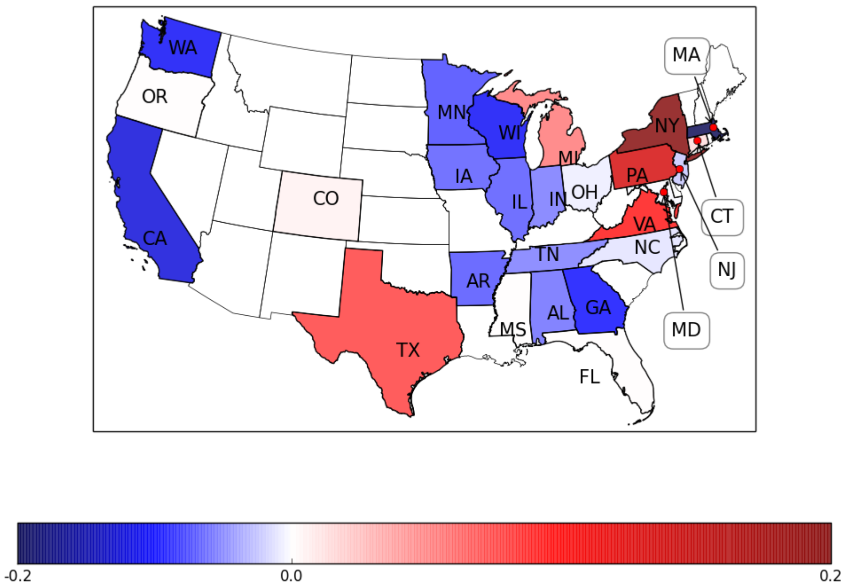
Figure J.2: Political distortion in young workers' welfare

in the opposite direction and causes a negative impact. Since the political economy has a decreased weight, it lowers infrastructure expenditures, and instead enhances the lump sum transfer. The cross-state impacts are sizeable, and vary within ±2% on an annual basis.



Note that colour codes indicate $\Delta\tilde{c}_{NB}^{y,k}$.

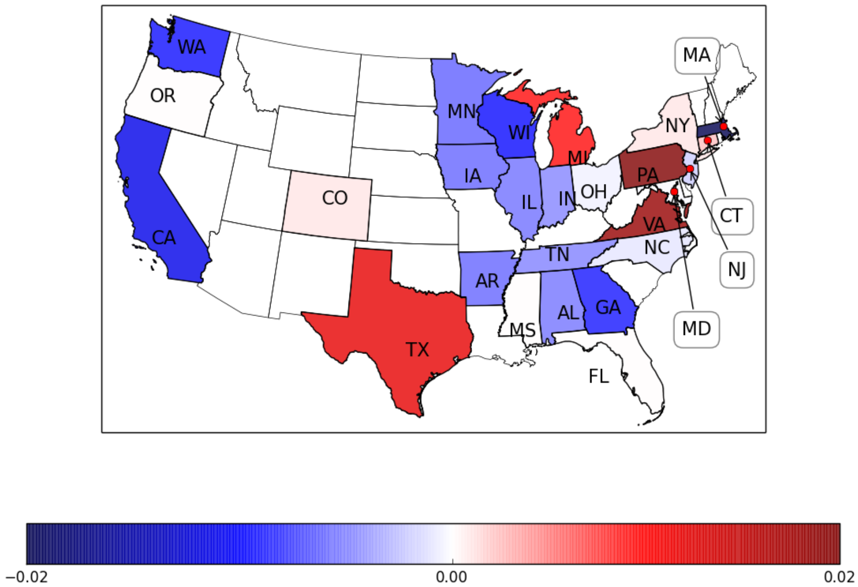
Figure J.3: Distortion in young capitalists' welfare due to political bias



Note that colour codes indicate $\Delta\tilde{c}_{NB}^{y,\ell}$.

Figure J.4: Distortion in young workers' welfare due to political bias

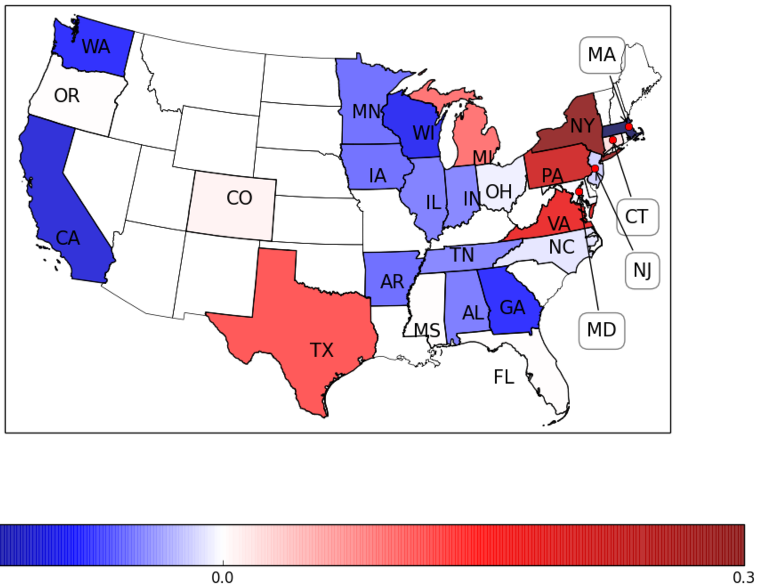
Next, we evaluate the welfare costs. Similarly to the evaluation of the political distortion, these costs are measured by $\Delta\tilde{c}_{NB}^{i,m} := \left(\tilde{c}^{i,m} - \tilde{c}_{NB}^{i,m}\right)/\tilde{c}_{NB}^{i,m}$ where $\tilde{c}_{NB}^{i,m} := c_{t,NB}^{i,m}/K_t$ is the consumption-capital ratio for i-aged, m-type individuals in the benchmark economy. For the same reason as in Section 4 in the paper, we only present



Note that the colour codes indicate $\Delta\nu_{NB}$ on an annual basis.

Figure K.1: Distortion in growth due to political bias

the welfare impacts on old agents; see Appendix J for the results for young agents. Figure K.2 shows the welfare cost to the capitalists. As is the case for political distortion in the main paper, the cross-state pattern of the heat map is identical to Figure K.1. Figure K.3 shows the welfare costs to old workers. In some states including Texas and Illinois, we observe opposite impacts, whereas in other states, including Virginia and Alabama, we observe parallel impacts. In the former case, the individual-level transfer to workers changes inversely with infrastructure expenditure. In the latter case, the transfer is shifted in the same direction as the infrastructure expenditure. For the same reason as in Section 4 of the main paper, the consequence of the distortion depends on the movement of agent fractions through the *ex-ante* choice.

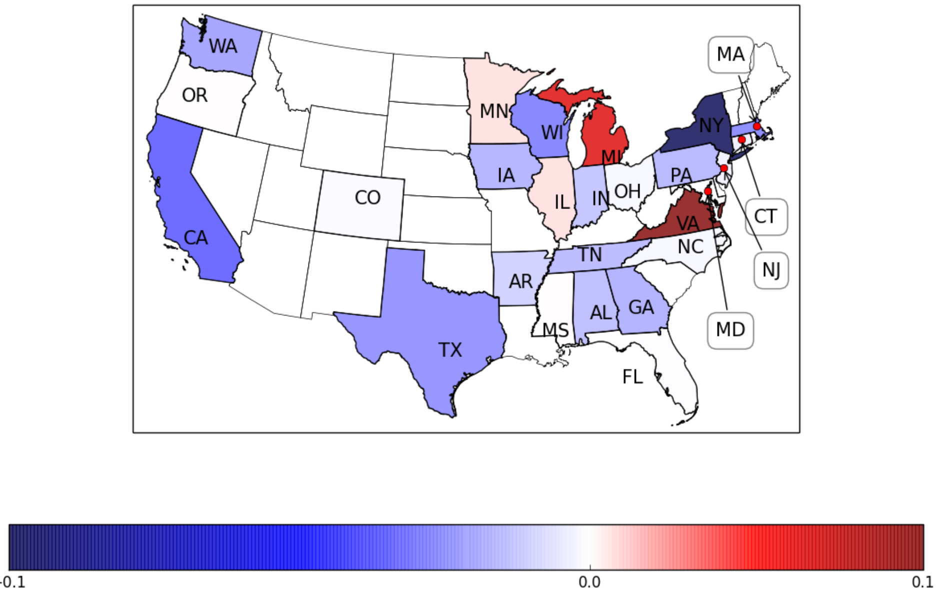


Note that the colour codes indicate $\Delta \tilde{c}_{NB}^{o,k}$.

Figure K.2: Distortion in old capitalists' welfare due to political bias

References

Acemoglu, Daron, *Introduction to Modern Economic Growth*, Princeton University Press, 2008.



Note that the colour codes indicate $\Delta \tilde{c}_{NB}^{o,\ell}$.

Figure K.3: Distortion in old workers' welfare due to political bias

Alesina, Alberto and Andrea Passalacqua, "The Political Economy of Government Debt," Working Paper 21821, National Bureau of Economic Research December 2015.

Austen-Smith, David and Jeffrey S. Banks, *Positive Political Theory II: Strategy and Structure*, The University of Michigan Press, 2005.

Barro, Robert J., "Government Spending in A Simple Model of Endogenous Growth," *Journal of Political Economy*, 1990, *98* (5), S103–S125.

Campbell, Albert Angus, Philip Ernest Converse, Warren E. Miller, and Donald Elkinton Stokes, *The American Voter*, Wiley & Sons, 1960.

Capello, Roberta and Peter Nijkamp, *Handbook of Regional Growth and Development Theories*, Edward Elgar Publishing, 2010.

Coe, Neil, Philip Kelly, and Henry W. C. Yeung, *Economic Geography: A Contemporary Introduction*, 2nd ed., Wiley-Blackwell, 2013.

Combes, Pierre-Philippe, Thierry Mayer, and Jacques-Franois Thisse, *Economic Geography: The Integration of Regions and Nations*, Princeton University Press, 2008.

Cooley, Thomas F and Edward C Prescott, "Economic Growth and Business Cycles," in "Frontiers of Business Cycle Research," Princeton University Press, 1995.

Domar, Evsey D., "Capital Expansion, Rate of Growth, and Employment," *Econometrica*, 1946, *14* (2), pp. 137–147.

Gerking, Shelby, "Measuring Productivity Growth in U.S. Regions: A Survey," *International Regional Science Review*, 1993, *16* (1-2), 155–185.

Golan, Amos, George G. Judge, and Douglas Miller, *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, Wiley, 1996.

Gomme, Paul and Peter Rupert, "Theory, Measurement and Calibration of Macroeconomic Models," *Journal of Monetary Economics*, 2007, *54* (2), 460–497.

Harris, Richard, "Models of Regional Growth: Past, Present and Future," *Journal of Economic Surveys*, 2011, *25* (5), 913–951.

Harrod, R.F., "An Essay in Dynamic Theory," *Economic Journal*, 1939, *49*, 14–33.

Heer, Burkhard and Alfred Maussner, *Dynamic General Equilibrium Modeling: Computational Methods and Applications*, Springer, 2009.

Hulten, Charles R. and Robert M. Schwab, "Regional Productivity Growth in U.S. Manufacturing: 1951-78," *The American Economic Review*, 1984, *74* (1), 152–162.

Hurd, M.D., "Mortality Risk and Bequests," *Econometrica*, 1989, *57* (4), 779–813.

Jaynes, E. T., "Information Theory and Statistical Mechanics," *Physical Review*, May 1957, *106* (4), 620–630.

_ , "Information Theory and Statistical Mechanics II," *Physical Review*, 1957, *108* (2), 171–190.

Jorgensen, Dale and Zvi Griliches, "The Explanation of Productivity Change," *Review of Economic Studies*, 1967, *34*, 249–280.

Leighley, Jan E. and Jonathan Nagler, *Who Votes Now?*, Princeton University Press, 2013.

Lindbeck, Assar and Jörgen Weibull, "Balanced-budget redistribution as the outcome of political competition," *Public Choice*, January 1987, *52* (3), 273–297.

Martin, R., "Institutional Approaches in Economic Geography," in Eric Sheppard and Trevor J. Barnes, eds., *A Companion to Economic Geography*, Wiley-Blackwell, 2002.

Mebane, Walter R. and Jonathan Wand, "Markov Chain Models for Rolling Cross-section Data: How Campain Events and Political Awareness Affect vote Intensions and Partisanship in the United States and Canada," in "1997 Annual Meeting of the Mid-west Political Science Association at Palmer House Hilton, Chicago, IL" 1997.

Mulligan, Casey B. and Xavier Sala-i-Martin, "Gerontocracy, Retirement, and Social Security," NBER Working Papers 7117, National Bureau of Economic Research, Inc May 1999.

North, Douglass C., *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, 1990.

Okabe, Tomohito, "Economic Growth, Politics and Institutions." PhD dissertation, Australian National University 2015.

Olson, Mancur, *The Rise and Decline of Nations*, New Haven: Yale University Press, 1984.

Persson, Torsten and Guido Tabellini, *Political Economics: Explaining Economic Policy*, MIT Press, 2002.

Rafiqui, Pernilla S., "Evolving Economic Landscapes: Why New Institutional Economics Matters for Economic Geography," *Journal of Economic Geography*, 2009, *9* (3), 329–353.

Razin, A., E. Sadka, and P. Swagel, "The Aging Population and the Size of the Welfare State," *Journal of Political Economy*, 2002, *110* (4), 900–918.

Ricketts, Lowell R. and Christopher J. Waller, "State and Local Debt: Growing Liabilities Jeopardize Fiscal Health," *Regional Economist*, October 2010.

Shannon, C.E., "A Mathematical Theory of Communication," *Bell Systems Technical Journal*, 1948, *27* (3), 379–423.

Solow, Robert M., "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 1957, *39*, 312–320.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt," *Econometrica*, 2012, *80* (6), 2785–2803.

Watts, Duncan, *Dictionary of American Government and Politics*, Edinburgh University Press, 2010.

Wolfinger, R.E. and S.J. Rosenstone, *Who Votes?*, Yale University Press, 1980.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, 2 ed., MIT Press, 2010.

Yamarik, Steven, "State-Level Capital and Investment: Updates and Implications," *Contemporary Economic Policy*, 2013, *31* (1), 62–72.