Appendix: For online publication

Contents

A State-level data on infrastructure

Table A.1 summarizes the state-level data on infrastructure.

State	Avg. Rank (Infra.)	# Good Group	Portion (Infra.)	Portion (GDP)	Avg. Rank (Estab.)	Portion (Estab.)
New York	1.708	24.000	0.072	0.081	2.439	0.072
California	1.833	24.000	0.071	0.133	1.000	0.114
Texas	2.458	24.000	0.071	0.079	2.659	0.069
Florida	4.000	24.000	0.064	0.050	4.000	0.060
Illinois	5.000	24.000	0.049	0.046	5.341	0.044
Ohio	6.542	24.000	0.035	0.036	7.000	0.039
New Jersey	7.125	24.000	0.034	0.034	8.415	0.033
Georgia	8.458	24.000	0.032	0.029	11.171	0.027
Pennsylvania	8.708	24.000	0.032	0.040	5.561	0.044
Massachusetts	9.708	24.000	0.030	0.027	12.341	0.024
Minnesota	10.458	24.000	0.029	0.018	18.561	0.019
North Carolina	12.208	24.000	0.025	0.027	9.976	0.028
Wisconsin	13.083	24.000	0.025	0.017	17.439	0.020
Washington	14.250	24.000	0.024	0.024	14.561	0.022
Virginia	14.458	24.000	0.024	0.027	12.585	0.025
Michigan	16.083	24.000	0.022	0.030	9.024	0.032
Tennessee	16.917	24.000	0.021	0.018	19.195	0.019
Missouri	18.167	24.000	0.019	0.018	15.171	0.021
Indiana	18.833	24.000	0.018	0.019	15.171	0.021
Kentucky	20.292	24.000	0.018	0.011	27.415	0.013
Louisiana	21.333	24.000	0.017	0.014	22.805	0.015
Iowa	21.625	24.000	0.017	0.010	28.951	0.012
Arizona	22.875	24.000	0.016	0.017	23.756	0.016
Colorado	25.625	15.000	0.015	0.017	19.439	0.018
Kansas	25.833	13.000	0.014	0.009	30.829	0.011
Alabama	26.000	23.000	0.015	0.012	24.951	0.014
Maryland	26.042	11.000	0.015	0.020	20.415	0.018
Connecticut	26.542	10.000	0.014	0.016	25.951	0.014
Oklahoma	29.458	0.000	0.012	0.010	27.634	0.013
Mississippi	30.208	0.000	0.011	0.006	33.317	0.009
Oregon	30.500	0.000	0.011	0.011	25.659	0.014
South Carolina	31.917	0.000	0.011	0.011	26.634	0.013
Nevada	33.083	0.000	0.010	0.008	38.000	0.006
Nebraska	34.417	0.000	0.010	0.006	34.927	0.007
Arkansas	34.708	0.000	0.010	0.007	32.439	0.009
New Mexico	35.542	0.000	0.010	0.006	37.000	0.006
West Virginia	37.000	0.000	0.009	0.004	37.244	0.006
Utah	38.375	0.000	0.008	0.007	34.122	0.008
Alaska	39.167	0.000	0.007	0.003	51.000	0.002
Hawaii	39.458	0.000	0.007	0.005	41.854	0.005
Idaho	41.667	0.000	0.006	0.004	40.512	0.005
Montana	41.958	0.000	0.006	0.002	42.512	0.004
Delaware	42.375	0.000	0.006	0.004	47.317	0.003
Wyoming	44.167	0.000	0.005	0.002	49.707	0.003
South Dakota	45.042	0.000	0.005	0.002	45.073	0.003
Rhode Island	46.083	0.000	0.004	0.003	42.963	0.004
Maine	47.208	0.000	0.004	0.004	39.098	0.005
North Dakota	47.500	0.000	0.004	0.002	47.146	0.003
New Hampshire	49.000	0.000	0.003	0.004	39.963	0.005
District of Columbia	50.000	0.000	0.002	0.007	47.927	0.003
Vermont	51.000	0.000	0.002	0.002	47.829	0.003

Table A.1: State-level summary

Notes: Avg. Rank (Infra.) is the average time-series ranking of infrastructure (this variable is the sorting variable). # Good Group is how many times the state belonged to the good infrastructure group (Max:24). Portion (Infra.) is the portion of infrastructure on average. Avg. Rank (Estab.) is the average time-series ranking of the number of establishments. Portion (Estab.) is the portion of establishments on average.

We outline the process for constructing state-level infrastructure capital as follows: We first compute the net investment on public and private capital stocks.

The state-level net public investment is approximated by the portion of aggregate net infrastructure investment, where the weight is obtained by the state-level real public highway infrastructure investment from Bennett, Kornfeld, Sichel, and Wasshausen (2020).[1] This is from the assumption that the infrastructure spending at the state level for each of the different items (e.g., highway, water supply, etc.) is identically distributed across the states.

The state-level net private investment is approximated by the portion of aggregate net non-residential fixed investment from National Income and Product Accounts (NIPA) data of Bureau of Economic Analysis (table 5.2.6), where the weight is obtained by the number of establishments at the state level from the Business Dynamics Statistics (BDS) at the United States Census Bureau. This approximation assumes that the establishment-level capital stock does not vary significantly across the establishments.[2]

After we obtain the net investment for public and private capital, we construct public and private capital stocks using the perpetual inventory method. For this approach, the initial capital stocks are needed for both public and private capital stocks. The state-level initial infrastructure stock is obtained by the portion of the infrastructure stock in 1977 from Bennett, Kornfeld, Sichel, and Wasshausen (2020), where the weight is from the highway infrastructure spending in 1977. The state-level initial private capital stock is from the portion of the aggregate private capital

[1]The aggregate net infrastructure investment is also from Bennett, Kornfeld, Sichel, and Wasshausen (2020). In the state-level calculation, the weight is computed in the following way:

$$weight_{it} = \frac{\text{highway infrastructure investment}_{it}}{\sum_i \text{highway infrastructure investment}_{it}}$$

[2]As a robustness check, we utilize the Census-based state-level private capital data in the manufacturing sector, which is available from Falk and Shelton (2018). This dataset provides the private capital k share in the good-infrastructure region as 0.83, which is close to our establishment-based value of 0.84.

stock in 1977 from NIPA of Bureau of Economic Analysis (table 4.1), where the weight is from the number of establishments in 1977. All the data is at the annual frequency. All real variables are chained in 2012 dollar value.

B Notes on the competitive equilibrium

We assume the optimal dividend payout policy fully internalizes the income tax of households, τ^h. Without this assumption, there would be an inefficient allocation of dividends, which is beyond the scope of this paper.[3] Firms earn tax benefit from tax shield out of the depreciated capital δk_t.

The value function of the household is as follows:

$$V_t(a_t) = \max_{c_t, a_{t+1}, L_t, B_{t+1}} log(c_t) - \sum_{j \in \{G,P\}} \omega_j \frac{\eta}{1+\frac{1}{\chi}} L_{jt}^{1+\frac{1}{\chi}} + \beta V_{t+1}(a_{t+1}) \tag{1}$$

$$\text{s.t. } c_t + \frac{a_{t+1}}{1+r_t} + \frac{B_{t+1}}{1+r_t^B} \tag{2}$$

$$= \sum_j \omega_j (w_{jt}\mathcal{E}_t + w_{jt}L_{jt})(1-\tau^h) + D_t(1-\tau^h) + (a_t - D_t) + T_t + B_t \tag{3}$$

A firms' value function is as follows:

$$J_t(z_t, k_t, j_t) = \max_{I_t, I_t^c} \quad \pi_t(z_t, k_t, j_t)(1-\tau^c)(1-\tau^h) + \delta\tau^c k_t(1-\tau^h) \tag{4}$$

$$+ \int_0^{\bar{\xi}} \max\{(-I_t - w_{jt}\xi - C(I_t, k_t))(1-\tau^h) + \frac{1}{1+r_t}\mathbb{E}J_{t+1}(z_{t+1}, k_{t+1}, j_{t+1}), \tag{5}$$

$$(-I_t^c - C(I_t^c, k))(1-\tau^h) + \frac{1}{1+r_t}\mathbb{E}J_{t+1}(z_{t+1}, k_{t+1}^c, j_{t+1})\} dG(\xi) \tag{6}$$

[3] Without this assumption, the firm's profit maximization would not take into account the household's income tax. This contrasts with the household's saving decision, which is based on the future after-income-tax dividend, leading to a distortionary effect of the corporate tax. Analyzing this distortionary effect is beyond the scope of this paper.

$$\text{s.t.} \qquad k_{t+1} = (1-\delta)k_t + I_t, \quad I_t \notin \Omega(k_t) = [-\nu k_t, \nu k_t] \tag{7}$$

$$k^c_{t+1} = (1-\delta)k_t + I^c_t, \quad I^c_t \in \Omega(k_t) \tag{8}$$

$$dG(\xi) = \frac{1}{\bar{\xi}} d\xi \tag{9}$$

$$\pi_t(z_t, k_t, j_t) = \max_{l_t} \; z_t x_{jt} \mathcal{C}(Y_{jt}) f(k_t, l_t, N_{jt}) - w_{jt} l_t \tag{10}$$

$$C(I_t, k_t) = \frac{\mu}{2} \left(\frac{I_t}{k_t}\right)^2 k_t \tag{11}$$

B.1 Interest rate and capital market

In the model, there are three competitive markets: the capital market and the regional labor markets. Thus, there are three prices to be determined endogenously. Define $p_t := U'(c_t) = 1/c_t$. Then, from the Euler equation of the representative household,

$$\beta \mathbb{E}_t \frac{U'(c_{t+1})}{U'(c_t)} = \frac{1}{1+r^B_t} \iff \beta \frac{p_{t+1}}{p_t} = \frac{1}{1+r^B_t} = \frac{1}{1+r_t} \quad (\because \text{no arbitrage}) \tag{12}$$

As there is no aggregate uncertainty, the expectation operator for the household can be lifted. Then, define a modified value function $\tilde{J}_t(z,k,j) = p_t J_t(z,k,j)$. We will show that the original recursive formulation can be rewritten with respect to the modified value function, which dispenses with the future price p_{t+1} in the formulation. Thus, p_t is the only price that needs to be traced.

The following is the original recursive formulation:

$$J_t(z_t, k_t, j_t) = \max_{I_t, I_t^c} \quad \pi_t(z_t, k_t, j_t)(1-\tau^c)(1-\tau^h) + \delta\tau^c k_t(1-\tau^h) \tag{13}$$

$$+ \int_0^{\bar{\xi}} \max\{(-I_t - w_{jt}\xi - C(I_t, k_t))(1-\tau^h) + \frac{1}{1+r_t}\mathbb{E}J_{t+1}(z_{t+1}, k_{t+1}, j_{t+1}), \tag{14}$$

$$(-I_t^c - C(I_t^c, k))(1-\tau^h) + \frac{1}{1+r_t}\mathbb{E}J_{t+1}(z_{t+1}, k_{t+1}^c, j_{t+1})\}dG(\xi) \tag{15}$$

We replace $\frac{1}{1+r_t}$ with $\beta\frac{p_{t+1}}{p_t}$. So we have,

$$J_t(z_t, k_t, j_t) = \max_{I_t, I_t^c} \quad \pi_t(z_t, k_t, j_t)(1-\tau^c)(1-\tau^h) + \delta\tau^c k_t(1-\tau^h) \tag{16}$$

$$+ \int_0^{\bar{\xi}} \max\{(-I_t - w_{jt}\xi - C(I_t, k_t))(1-\tau^h) + \beta\frac{p_{t+1}}{p_t}\mathbb{E}J_{t+1}(z_{t+1}, k_{t+1}, j_{t+1}), \tag{17}$$

$$(-I_t^c - C(I_t^c, k))(1-\tau^h) + \beta\frac{p_{t+1}}{p_t}\mathbb{E}J_{t+1}(z_{t+1}, k_{t+1}^c, j_{t+1})\}dG(\xi) \tag{18}$$

Then, multiply p_t to both sides. It leads to

$$p_t J_t(z_t, k_t, j_t) = \max_{I_t, I_t^c} \quad p_t\pi_t(z_t, k_t, j_t)(1-\tau^c)(1-\tau^h) + p_t\delta\tau^c k_t(1-\tau^h) \tag{19}$$

$$+ \int_0^{\bar{\xi}} \max\{(-p_t I_t - p_t w_{jt}\xi - p_t C(I_t, k_t))(1-\tau^h) + \beta\mathbb{E}p_{t+1}J_{t+1}(z_{t+1}, k_{t+1}, j_{t+1}), \tag{20}$$

$$(-I_t^c - C(I_t^c, k))(1-\tau^h) + \beta\mathbb{E}p_{t+1}J_{t+1}(z_{t+1}, k_{t+1}^c, j_{t+1})\}dG(\xi) \tag{21}$$

Thus, we have

$$\widetilde{J}_t(z_t, k_t, j_t) = \max_{I_t, I_t^c} \quad p_t \pi_t(z_t, k_t, j_t)(1-\tau^c)(1-\tau^h) + p_t \delta \tau^c k_t (1-\tau^h) \tag{22}$$

$$+ \int_0^{\overline{\xi}} \max\{(-p_t I_t - p_t w_{jt} \xi - p_t C(I_t, k_t))(1-\tau^h) + \beta \mathbb{E} \widetilde{J}_{t+1}(z_{t+1}, k_{t+1}, j_{t+1}), \tag{23}$$

$$(-I_t^c - C(I_t^c, k))(1-\tau^h) + \beta \mathbb{E} \widetilde{J}_{t+1}(z_{t+1}, k_{t+1}^c, j_{t+1})\} dG(\xi) \tag{24}$$

Therefore, a firm's inter-temporal decision is perfectly characterized by tracing the contemporaneous price

$$p_t = U'(c_t) = 1/c_t.$$

B.2 Wage and labor market

From the representative household's intra-temporal optimality condition (with respect to the labor supply),

$$\omega_j \eta L_{jt}^{\frac{1}{\chi}} = U'(c_t) \omega_j w_{jt} (1-\tau^h) \quad \text{for } \forall j \in \{G, B\} \tag{25}$$

Therefore,

$$\eta L_{jt}^{\frac{1}{\chi}} = p_t w_{jt} (1-\tau^h) \implies w_{jt} = \frac{\eta}{p_t (1-\tau^h)} L_{jt}^{\frac{1}{\chi}} \tag{26}$$

The optimal labor supply L_{jt} depends upon w_{jt}, and w_{jt} can be determined only when the labor supply L_{jt} is known, leading to a fixed-point problem. Therefore,

w_{jt} needs to be tracked together with p_t for the computation.[4]

C Estimation details

C.1 Challenges of estimating a general equilibrium model with the existing SMM

The fundamental objective of the SMM method is to minimize the discrepancy between the moments generated by the model and those observed empirically. Let Θ denote the parameters of interest and $\hat{\mathbf{m}}$ denote the vector of M moments from the data for estimation. Under SMM, the moment conditions to satisfy are $\hat{\mathbf{m}} - \mathbf{m}(\Theta) = \mathbf{0}$, where $\mathbf{m}(\Theta)$ is the model's prediction for the moments under parameter Θ and $\mathbf{0}$ is a zero vector of length M.

Suppose we estimate parameters of the model in which market clearing conditions need to be satisfied as general equilibrium conditions. Given each candidate parameter vector, the model is solved with an additional loop that makes sure the market clearing conditions become zero with numerical precision. This additional layer regarding general equilibrium conditions is likely to result in prohibitively high computational costs.

C.2 Implementation of SMM in a Bayesian way

The limited-information Bayesian method, as described in Kim (2002) and later advocated by Christiano, Trabandt, and Walentin (2010) and Fernández-Villaverde, Rubio-Ramírez, and Schorfheide (2016) among others, can be viewed as the Bayesian version of the simulated method of moments (SMM). The limited-information Bayesian method only uses a set of moments from the data for parameter inference.

[4] If $\chi \to \infty$, p_t is the only price to be tracked as in Khan and Thomas (2008).

Let Θ denote the parameters of interest and $\hat{\mathbf{m}}$ denote the vector of M empirical moments from the data for estimation. The likelihood of $\hat{\mathbf{m}}$ conditional on Θ is approximately

$$f(\hat{\mathbf{m}}|\Theta) = (2\pi)^{-\frac{M}{2}} \exp\left[-\frac{1}{2}\left(\hat{\mathbf{m}} - \mathbf{m}(\Theta)\right)'\left(\hat{\mathbf{m}} - \mathbf{m}(\Theta)\right)\right],$$

where $\mathbf{m}(\Theta)$ is the model's prediction for the moments under parameter Θ.[5] Bayes' theorem tells us that the posterior density $f(\Theta|\hat{\mathbf{m}})$ is proportional to the product of the likelihood $f(\hat{\mathbf{m}}|\Theta)$ and the prior density $p(\Theta)$:

$$f(\Theta|\hat{\mathbf{m}}) \propto f(\hat{\mathbf{m}}|\Theta)p(\Theta),$$

and we can then apply the standard Markov Chain Monte Carlo (MCMC) techniques to obtain a sequence of random samples from the posterior distribution.

C.3 Implementation of multiple-block Metropolis-Hastings

We use the multiple-block Metropolis Hastings algorithm to estimate the model parameters as well as finding market clearing prices. Let's denote the moments to match (including the market clearing conditions) as $y \equiv [\hat{\mathbf{m}}, \mathbf{0}]$. $\hat{\mathbf{m}}$ is for the moments constructed from the data and $\mathbf{0}$ is associated with solving for general equilibrium. We break the parameter space into two blocks as follows: $\Theta = (\Theta^1, \Theta^2)$ where Θ^1 is for the price block and Θ^2 is for the other model parameter block. Starting from an initial value $\Theta_0 = (\Theta_0^1, \Theta_0^2)$, the algorithm works as follows:
For iteration $j = 1, \ldots, M$, and for block $k = 1, 2$.

[5] Each moment's discrepancy is assumed to contribute equally. Alternatively, we can use the consistent estimator for the covariance matrix of $\hat{\mathbf{m}}$ and its inverse as the weighting matrix. This alternative weighting scheme has little impact on our estimation results.

- Propose a value $\tilde{\Theta}^k$ for the kth block, conditional on Θ^k_{j-1} for the kth block and the current value of the other block (Θ^{-k}). Θ^{-k} stands for the remaining block except for the kth block.[6]

- Compute the acceptance probability $\alpha^k = \min\left\{1, \frac{f(\tilde{\Theta}^k|\Theta^{-k},y)}{f(\Theta^k_{j-1}|\Theta^{-k},y)}\right\}$.
  Update the kth block as

$$\Theta^k_j = \begin{cases} \tilde{\Theta}^k & \text{w.p. } \alpha^k \\ \Theta^k_{j-1} & \text{w.p. } (1-\alpha^k) \end{cases}$$

For each iteration, we first update the price block conditional on the previous iteration's value for the price block and the remaining model parameter block. Then we sequentially update the model parameter block conditional on the updated price block.[7]

We apply the multiple-block RWMH algorithm to simulate draws from the posterior density $f(\Theta|\hat{\mathbf{m}})$ with uniform priors. The posterior distribution is characterized by a sequence of 3000 draws. We initialize the chain at the point estimate from particle swam optimization routine from MATLAB.

C.4 Externally calibrated parameters

Externally calibrated parameters are reported in Table C.2. We set $\beta = 0.96$, $\alpha = 0.28$, and $\gamma = 0.64$, following common values in the literature. For the average

[6] In our application, $\Theta^{-1} = \Theta^2$ and $\Theta^{-2} = \Theta^1$.

[7] For the accuracy boost in the price block, we consider an extra step of the typical general equilibrium routine. To be specific, given Θ^1, we can compute the implied level of prices Θ^{1*} based on the market clearing conditions. When $\Theta^1 = \Theta^{1*}$, the GE is obtained associated with Θ^2. The extra loop iteratively updates Θ^1 based on convex combination with corresponding Θ^{1*}, while Θ^2 is fixed. The convergence is achieved due to the *stability* of the general equilibrium. However, this step is optional for the method's application.

β	discount rate	0.96	φ	infrastructure spending	0.09
α	capital share	0.28	s	time to build	1.00
γ	labor share	0.64	χ	Frisch elasticity	4.00
τ^h	household income tax rate	0.15	δ	private capital depreciation	0.09
τ^c	corporate tax rate	0.27	$\delta_{\mathcal{N}}$	public capital depreciation	0.02
$\mathcal{E}$	public employment	0.05	ρ_z	z shock persistence	0.75
ζ_G	Region G's infra. portion	0.81	σ_z	z shock volatility	0.13
ω_G	Region G's labor force share	0.73	ι	congestion elasticity	0.01

Table C.2: Externally calibrated parameters

Notes: Each period in the model corresponds to one year in the data.

of household income tax rate, we use 0.15 as in Krueger and Wu (2021) where they compute the tax rate with the data from Blundell, Pistaferri, and Saporta-Eksten (2016). For corporate tax rate, we use 0.27 from Gravelle (2014) that is the effective tax paid after deductions and credits. We use 0.05 for the fraction of public employment, using the FRED data on the government employees ($USGOVT$) and the private employees ($USPRIV$). We use 0.09 for the infrastructure spending out of tax revenue. This comes from the fact that the infrastructure spending as share of GDP is 2.4% and the tax revenue as share of GDP is 27.1%. We assume one year of time-to-build for the baseline analysis. We set Frisch elasticity to be 4 as in Ramey (2020). We use 0.09 for the private capital depreciation rate, and 0.02 for the public capital depreciation rate from the BEA depreciation data. We use the estimates of the persistence and volatility of the idiosyncratic productivity shocks in Lee (2025), which applies the methodology of Ackerberg, Caves, and Frazer (2015) on Compustat data from Standard and Poor's. The congestion elasticity parameter is set at the level where the erosion of the steady-state aggregate output due to the congestion effect is at 3%.[8]

[8]We provide a robustness check in Appendix G for different levels of congestion effects ranging from 1% to 10%. Our main results stay unaffected over the different choices of the congestion parameter.

C.5 State grouping

Our model captures state-level variations by defining two regions, P and G, based on the infrastructure size and geographical proximity, as illustrated in Figure C.1. The brown areas represent poor-infrastructure regions, primarily in the West, while the green areas indicate good-infrastructure regions. Note that California, Texas, and Washington are excluded from the West, as they do not fit the poor-infrastructure category. Similarly, poor-infrastructure states in the East are also omitted. Additionally, transitional states—such as Minnesota, Iowa, Missouri, and Louisiana—are excluded to minimize spillover effects. These excluded states, shown in grey, are not considered in the calculation of state-level infrastructure and private capital shares.[9]

Our classification reassigns Texas and California to the good-infrastructure group, ensuring a more representative allocation.[10] Since their neighboring states exhibit minimal spillover effects and lack major MSAs near their borders, this grouping better reflects infrastructure disparities. Table 2 in the main text presents summary statistics comparing poor- and good-infrastructure regions, with data sourced from Bennett, Kornfeld, Sichel, and Wasshausen (2020).[11]

The transition probabilities are set to be persistent ($\pi_{PP} = 0.90, \pi_{GG} = 0.98$).[12] The infrastructure portion for group G, ζ_G, is set at 0.81, and the Poor's portion ζ_P is 0.19.[13]

[9]We thank an anonymous referee for suggesting the geographical grouping. In our working paper, regions were classified using the median ranking of per capita infrastructure capital stock from 1994 to 2017.

[10]Further state-level details are in Appendix A.

[11]Details on the state-level data sources and variable construction are presented in Appendix A.

[12]Transition probabilities are constructed using the state-level data in Table A.1 in Appendix A. Specifically, we use the two moments: 1) the transition probability from the Good to Poor region and 2) the ratio of the number of firms between the two regions. We check the robustness of our main results over a different specification of the transition probability in Appendix G.

[13]If we standardize the infrastructure capital stock of the poor and good groups by their respec-

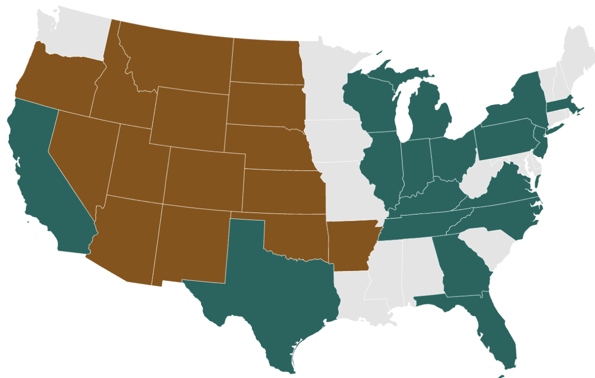


Figure C.1: Regions with good vs. poor infrastructure

C.6 Targeted moments and sensitivity analysis

Following Cooper and Haltiwanger (2006), we assess the sensitivity of the selected moment targets to illustrate the usefulness of incorporating the high region's private capital share in inferring λ. Table C.3 reports the moment sensitivity when the parameter changes by 1%. The number is obtained by the average absolute log changes for 1% increase and 1% decrease.

Target moment	θ	λ	G	x
Private-to-infrastructure capital ratio	1.782	0.564	0.1400	0.165
High region's private capital k portion	0.044	0.015	0.002	0.012
Government spending to output ratio	5.553	1.659	1.046	0.229
High region's output y portion	0.022	0.013	0.003	0.001

Table C.3: Selected moment sensitivities in structural parameter variations

The private-to-infrastructure capital ratio is most sensitive to variations in θ, followed by λ. After accounting for θ, which has a stronger association with the private-to-infrastructure capital ratio, λ emerges as the second most influential parameter in disciplining the good-infrastructure region's share of private capital k.

tive population sizes, we find that $\zeta_G = 0.84$, which is close to the value of 0.81 used in our analysis.

Consequently, we suggest that the high region's private capital k portion serves as a useful moment for refining our understanding of λ.

C.7 External validation with empirical state-level elasticity

As external validation, we compute the state-level elasticity from our model and compare it to the empirical estimate using the state-level data. We find that the state-level input elasticity from our model indicates the complementarity between private and public capital and this is consistent with the empirical elasticity obtained from the state-level production function estimation.

C.7.1 State-level elasticity from the model

In our model, the infrastructure stock is shared among the firms in the same region. We conduct the state-level aggregation as follows: we fix the firm-level estimates except for the elasticity λ and spatial productivity heterogeneity x_1.[14] We estimate these two parameters under the state-level production functions.[15]

$$\begin{bmatrix} x_1 \left(\theta k_1^{\frac{\lambda-1}{\lambda}} + (1-\theta) N_1^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}\alpha} l_1^{\gamma} \\ \left(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta) N_2^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}\alpha} l_2^{\gamma} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{27}$$

where (x_1, λ) are unknown, while all the other allocations and parameters, $(y_1, y_2, k_1, k_2, N_1, N_2, l_1, l_2, \theta, \alpha, \gamma)$ are obtained from the estimated baseline model.[16]

[14]Since the production function in our model is decreasing returns to scale, there is no guarantee that the firm-level elasticity and productivity is aggregated to have the same value in the state-level.

[15]We cannot identify the public capital stock share, θ separately from the elasticity, λ in the state-level model. This is the main reason why we introduce the micro-level heterogeneity in our structural model. Therefore, in the state-level model, we fix the public capital stock share at the firm-level estimate.

[16]The two parameters (x_1, λ) are obtained from the exact identification.

Using the same nonlinear least squares (NLLS) optimization used in An, Kangur, and Papageorgiou (2019), we get the estimate of the state-level production function $(x_1, \lambda) = (1.766, 0.349)$.

If we assume a CRS state-level production function with $\gamma = 1 - \alpha$, the estimates are $(x_1, \lambda) = (1.923, 0.482)$. Therefore, our model suggests that public capital and private capital are gross complements at the state level.[17]

It is important to emphasize that we compare the NLLS results derived from model-simulated data with those from empirical data (as in the following section), serving as an external validation. Our model validation aligns with the concept of indirect inference, which involves comparing the conditional correlation obtained from OLS regression using simulated data with that from empirical data.

C.7.2 State-level elasticity from the data

Using the state-level data, we estimate the elasticity of substitution between private and public capital given a CES production technology. We closely follow An, Kangur, and Papageorgiou (2019) in which the elasticity is estimated using the nonlinear least squares using the following:

$$\ln\left(\frac{Y_{it}}{Y_{i,t-1}}\right) = c + (1-a)\ln\left(\frac{L_{it}}{L_{i,t-1}}\right) + \frac{a}{\psi}\ln\left[\frac{bK_{it}^{\psi} + (1-b)N_{it}^{\psi}}{bK_{i,t-1}^{\psi} + (1-b)N_{i,t-1}^{\psi}}\right] + (\epsilon_{it} - \epsilon_{i,t-1}).$$

i denotes the state, t denotes the time, and ϵ is the error term. Y is the output, K is the private capital stock, N is the infrastructure capital stock, and L is employment. ψ is the capital substitution parameter which implies a public-private capital elasticity of substitution (ES) of $1/(1-\psi)$.

Table C.4 shows the estimation results from nonlinear least squares. The elas-

[17] Among the unreported results, we estimate the Cobb-Douglas production function as in Baxter and King (1993) using the simulated state-level data. The returns to scale parameter for the public capital is estimated to be greater than 0, consistent with the increasing returns to scale. The result is available upon request.

	Estimates	90% confidence interval
a	0.402	[0.351, 0.453]
b	0.070	[0.018, 0.123]
Elasticity of substitution	0.445	[-0.099, 0.989]

Table C.4: Results from nonlinear least squares estimation

Notes: Elasticity of substitution is $\frac{1}{(1-\psi)}$. Its confidence interval is derived by the delta method.

ticity of substitution between public and private capital is estimated to be 0.445.[18] In other words, the state-level variations indicate the complementarity between private and public capital. However, this result does not imply the complementarity between private and public capital at the firm level. In fact, the private and public capitals can be gross substitutes at the firm level, whereas they are gross complements at the state level.

It is worth noting that our model bridges the gap between the firm-level estimates and the state-level estimates. According to our estimates, private capital is a gross substitute for public capital at the firm level, while it is a gross complement of public capital at the state level. At the state level, the elasticity of substitution includes a good public nature of the infrastructure benefiting all firms in a state. Therefore, the non-rivalry of the infrastructure generates the complementarity between the state-level private capital and the public capital.

[18] As robustness check, we apply GMM estimation where $L_{it-2}, K_{it-2}, N_{it-2}$ are used in exogeneity conditions. The elasticity of substitution is estimated to be 0.44. This result is available upon request.

D Additional quantitative analysis

D.1 Efficiency vs. Equality: Cross-state analysis

In this section, we analyze the cross-state heterogeneous impact of the infrastructure investment. In our structural analysis, we divide all the states in the U.S. into two groups: Poor vs. Good by geographical proximity while controlling for potential spill-over effects.[19]

Table D.5: Cross-state inequality in short-run fiscal multipliers

	Baseline			Partial eq.			Equal spending		
	Total	Poor	Good	Total	Poor	Good	Total	Poor	Good
Y	1.1494	0.0537	1.0956	1.4142	0.0605	1.3537	0.7489	0.1468	0.6021
Inv.	-0.2055	-0.0148	-0.1907	0.1454	0.0087	0.1367	-0.2200	-0.0120	-0.2080
Earnings	2.9970	0.1653	2.8318	3.3815	0.1800	3.2015	1.2877	0.1775	1.1101
C	2.5771	0.0560	2.5212				0.9483	0.0598	0.8885

Table D.5 reports the heterogeneous state-specific fiscal multipliers. Per \$1 spending, out of the total output increase of \$1.149, \$1.096 goes to the output increase in the Good states, while only the incremental of \$0.054 belongs to the Poor states. In contrast, the crowding-out effect of the private investment is significantly greater in the Good states, featuring -0.191 out of total crowding-out effect of -0.206. In terms of the earnings, out of the total increase of \$2.997 per \$1 fiscal spending, \$0.165 belongs to the Poor states, while \$2.832 goes to the Good states. The aggravation of the cross-state inequality is more severe in the consumption side, as the infrastructure spending is financed by the lumpy-sum taxation impartially across the states, while the benefit of the spending is relatively more skewed

[19]The assumptions necessary for the cross-state analysis are specified in Appendix E.

to the Good states.[20]

In the baseline model, the general equilibrium effect weakly alleviates the unequal distribution of the benefit of the fiscal spending across the states, as can be seen from the earnings ratio between the two states: the earnings multiplier ratio is smaller than the counterpart in the partial equilibrium $(2.832/0.165 < 3.202/0.180)$. This is mainly due to the Good state's substantially more susceptible investment response to the GE crowding-out effect than the Poor state's investment (from 0.137 to -0.191 vs. from 0.009 to -0.015).

Table D.6: Region-specific taxes and short-run fiscal multipliers

	Baseline			Taxing Good region			Taxing Poor region		
	Total	Poor	Good	Total	Poor	Good	Total	Poor	Good
Y	1.1494	0.0537	1.0956	1.0036	0.0501	0.9535	1.2994	0.0575	1.2419
Inv.	-0.2055	-0.0148	-0.1907	-0.3777	-0.0257	-0.3519	-0.0306	-0.0036	-0.0270
Earnings	2.9970	0.1653	2.8318	0.3009	0.1074	0.1935	5.7270	0.2247	5.5023
C	2.5771	0.0560	2.5212	0.5556	1.0159	-0.4603	4.6214	-0.9030	5.5244

Then, we consider a counterfactual policy experiment where the fiscal shock is equally spent on the Good and Poor states. The outcome of this experiment shows a drastic trade-off a government would face between aggregate-level efficiency and cross-state equality. Compared to the status quo policy, equal spending leads to a substantially lower output multiplier (0.749) and a greater degree of crowding-out effect. Not surprisingly, the Good states' fiscal multiplier drops down significantly (0.602 vs. 1.096). However, the Poor states' output fiscal multiplier is around three times greater than the baseline (0.054 vs. 0.147); the earnings multiplier is around 8% greater than the baseline (0.165 vs. 0.178). This result is

[20]The tax is imposed more lightly for the Poor states than the Good states, but the tax burden is significantly greater than the benefit in the poor states in the baseline model.

largely due to the estimated state-specific productivity differences across the states. Also, due to the non-rivalry, the cross-state heterogeneity in the number of firms (capital) utilizing the additional infrastructure strongly affects the policy outcome.

The next policy experiment is counterfactual tax financing for fiscal spending shock. The middle and right sections of Table D.6 report the fiscal multipliers when the shock is purely financed by levying an additional lump-sum tax solely on the Good and Poor regions, respectively. The key channel of this experiment is the wealth effect, where reduced consumption affects the regional labor supply and the wage. With the Good region taxed, the aggregate fiscal multiplier decreases (from 1.149 to 1.004), while it increases for the other scenario (from 1.149 to 1.299). These results are crucially driven by the different degrees of crowding-out effects through consumption channels. The results also show a stark trade-off between efficiency and inequality across the states driven by different tax schemes. A greater aggregate-level efficiency is achieved once the tax is levied on the Poor region with a significantly heightened fiscal multiplier (from 1.149 to 1.299), while the cross-state inequality substantially worsens despite positive output gains for both states. Especially when it comes to consumption, inequality is substantially aggravated due to the direct tax effect, which supports the boost in aggregate output.

As this paper does not focus on the normative optimal policy, the implication is limited to a positive evaluation, but the evaluated trade-offs between efficiency and equality are highly policy-relevant takeaways of this model.

D.2 Fiscal multipliers and corporate taxation

In this section, we compare the fiscal multipliers when the infrastructure spending is combined with different tax policies. Three different policies are considered. The first policy is decreasing the corporate tax rate by 33% from the baseline level (27%→18%) for the period of shock and the following year (2 years). The second policy uses the baseline level (27%), and the last policy increases the corporate tax rate by 33% from the baseline level (27%→36%) for the initial two post-shock years.[21] The remaining balance in the fiscal budget after the change in taxation is financed by the lump-sum tax. Thus, the third policy collects the least amount of lump-sum tax among the three policies.[22]

Table D.7 reports the fiscal multipliers across the three corporate tax policies. In the first policy with low corporate tax, the short-run multiplier is around 1.211, which is the greatest among the three. In the last policy with high corporate tax, the short-run multiplier is around 1.087, which is the lowest among the three. The same ranking is observed for the long-run multipliers.

One of the main channels that cause the differences in the fiscal multipliers is the firm-level investment. When the fiscal spending is combined with the low corporate tax policy, due to the increased incentive of cumulating the future capital stock, the private investment significantly less crowds out. A similar pattern is observed in the long-run fiscal multipliers of private investment.

[21] The third policy mimics the Biden administration's original plan to increase the corporate tax rate by 33%. As our baseline tax level is 27% while the corporate tax rate of 2022 is 21%, there is a level difference in the tax rate.

[22] Our fiscal multiplier analysis is based on the impulse response to the MIT fiscal spending shock under perfect foresight. Therefore, the representative household becomes indifferent between lump-sum tax financing and debt financing as long as the lifetime income is unaffected. If the model considers household heterogeneity under the borrowing limit and frictional financial market, this indifference collapses, leading to divergent fiscal multipliers between tax financing and debt financing as in Hagedorn, Manovskii, and Mitman (2019).

Table D.7: Fiscal multipliers

Fiscal multipliers	Low Tax	Baseline	High Tax
Output			
Short-run	1.2112	1.1494	1.0867
Long-run	2.8731	2.6414	2.4084
Short-run (2 years)			
Consumption	2.4789	2.5771	2.6721
Investment	-0.1528	-0.2055	-0.2585
Public capital	1.6990	1.7039	1.7088
Labor income	3.0097	2.9970	2.9813
Long-run (5 years)			
Consumption	9.2564	8.7296	8.2003
Investment	-0.1033	-0.1782	-0.2533
Public capital	4.0988	4.1034	4.1080
Labor income	10.1780	9.4841	8.7883

The differences in the response of private capital investment to the fiscal policy lead to the differences in the labor income response. The greater the private investment, the greater the employment effect on the economy. In the low corporate tax policy, the labor income multiplier is 3.010; in the baseline corporate tax policy, the labor income multiplier is 2.997; in the high corporate tax policy, the labor income multiplier is 2.981.

However, the low corporate tax policy is not a free lunch. The low corporate tax policy leads to the lowest consumption multiplier of 2.479 in the short run. This is because this tax policy requires the greatest lump-sum tax to finance the spending shock. This clearly shows what is the trade-offs in corporate tax policies; the low tax policy sacrifices the short-run welfare to achieve long-run welfare. In the long run, due to the private investment and labor income channels, the fiscal multiplier is the greatest for the low corporate tax policy.

D.3 The role of time to build

In this section, we analyze the role of time to build on the fiscal multiplier. On top of the one-year time to build in the baseline, we assume there is an extra year of time to build for capital stock to be utilizable after the investment as in Ramey (2020) (two years, in total). Therefore, the law of motion of the public capital stock is as follows:[23]

$$\mathcal{N}_{A,t+2} = \mathcal{N}_{A,t+1}(1-\delta_{\mathcal{N}}) + \mathcal{F}_t - \frac{\mu}{2}\left(\frac{\mathcal{F}_t}{\mathcal{N}_{A,t+1}}\right)^2 \mathcal{N}_{A,t+1} \tag{28}$$

$$\mathcal{N}_{j,t+1} = \zeta_j \mathcal{N}_{A,t+1} \quad \text{for } j \in \{P, G\} \tag{29}$$

$$F_t = \begin{cases} F^{ss} + \Delta G & \text{if } t = 1 \\ F^{ss} & \text{otherwise} \end{cases} \tag{30}$$

where F^{ss} is the stationary equilibrium level of fiscal spending on infrastructure. Due to the time lag between the fiscal policy shock and the arrival of the public capital stock, there exists a news component in the policy, which will be analyzed further in this section.

For this analysis, the fiscal multiplier is measured by the sum of the present values over the first three years for the short run and over the six years for the long run after the initial fiscal spending shock.[24]

Table D.8 reports the fiscal multipliers when there is time to build of two years. The first column is the general equilibrium multipliers under the two years of the time to build; the second reports the same one in the partial equilibrium; the third is the baseline model; and the last is the baseline model in the partial equilibrium. The output fiscal multiplier decreases in general equilibrium, when the time-to-build is extended to two years (1.1494→0.6231), consistent with Ramey (2020). On

[23] For the consistency in the notation with the previous formulations, we leave the time index of the future public capital stock to be $t+1+s$ where $s=1$.

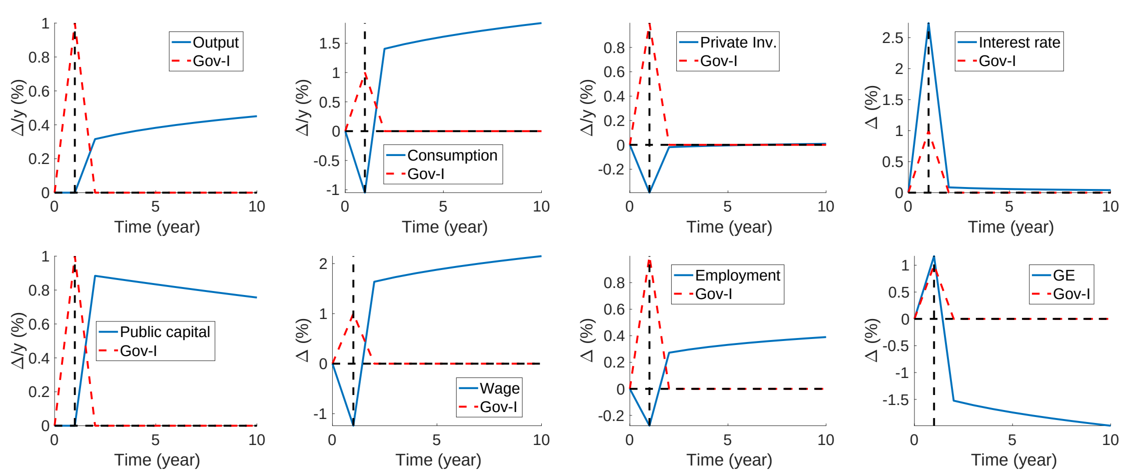
[24] Previously, it was 2 years for the short run and 5 years for the long run without the extended time to build.

Table D.8: Fiscal multipliers across the states under time to build of two years

Fiscal multipliers	T2B	T2B - PE	Baseline	Baseline - PE
Output				
Short-run	0.6231	1.5191	1.1494	1.4142
Long-run	1.7220	4.0969	2.6414	3.8364
Short-run (2 years)				
Consumption	1.6859		2.5771	
Investment	-0.4230	0.2227	-0.2055	0.1454
Public capital	1.6574	1.6807	1.7039	1.7161
Labor income	1.6867	2.6197	2.9970	3.3815
Long-run (5 years)				
Consumption	6.3164		8.7296	
Investment	-0.4364	0.4508	-0.1782	0.3669
Public capital	4.0568	4.0817	4.1034	4.1170
Labor income	6.5475	8.5200	9.4841	10.5524

the other hand, the output fiscal multiplier increases in the time to build in the partial equilibrium (1.4142→1.5191).

Figure D.2: The impulse responses to the infrastructure spending shock



To illustrate the role of the extended time to build, Figure D.2 plots the impulse responses of equilibrium allocations. Due to the extended time to build, the public capital spikes one year after the fiscal spending shock. As the fiscal spending shock hits, consumption immediately drops, as the lump-sum tax immediately puts downward pressure on the household's consumption. On the other hand, the production side does not face any direct change in the infrastructure until one year after the shock, only facing interest increases due to the consumption drop. Therefore, the private investment significantly decreases before the public capital becomes available. In the following period ($t = 2$), the added public capital arrives, so the production increases with private capital substantially shrunk due to the previous period's crowding-out effect. The increased output leads to an increase in employment and wages.

The news effect impacts the fiscal multiplier in the partial equilibrium, as it allows the agents with the rational expectation to adjust their allocations optimally even before the spending shock is capitalized. The news effect is the key channel boosting the fiscal multipliers in the partial equilibrium. However, this effect is dominated by changes in the price once we consider the general equilibrium effect. The agents' adjustment before the shock capitalization is followed by interest rate adjustment, dampening the fiscal multiplier even in a greater magnitude than the one-year time-to-build. This is because the interest rate adjustment occurs at one time, and the increased cost of investment in the period before the spending shock leads to a lowered capital stock. Under real friction, such as the convex adjustment cost, the lowered capital stock leads to a greater adjustment cost in the following period when the fiscal spending shock materializes, leading to a substantially dampened fiscal multiplier. Therefore, this is an outcome of the interaction between the news effect and the real friction.

D.4 The marginal product of private and infrastructure capital

In this section, we assess the equilibrium level of the marginal product of private and public capital stock. Table D.9 shows the marginal product of private and infrastructure capital stocks for the entire economy (column 1), the Good state (column 2), and the Poor state (column 3). In this economy, due to the presence of the capital adjustment cost at the firm level, the marginal product of capital varies across the firms.[25] We use the average marginal product of capital for the analysis.

Table D.9: The marginal product of private and public capital

Marginal product of capital (MPK)	Aggregate	Good state	Poor state
Private	0.2838	0.3339	0.0444
Infrastructure	0.3654	0.4074	0.1644

The marginal product of public capital stock is substantially higher than the private counterpart. This shows that the current stock of public capital is less than the socially desired level.[26] Moreover, the public-to-private MPK ratio is more than twice greater in the Poor state than in the Good state. This shows that the relative shortage of public capital provision is more severe in the Poor state than in the Good state in equilibrium. However, this relative shortage is an efficient outcome in the model, so the normative interpretation is limited.

[25] The convex adjustment cost depends on the capital stock of the firm, which makes the marginal cost of investment different across the firms. This leads to the heterogeneous marginal product of capital stock in equilibrium.

[26] If there were a competitive market for public capital, the price of the public capital would adjust in the direction to equate the shadow value of private and public capital.

D.5 Changes in equilibrium allocations due to the fiscal shock

Table D.10 presents changes in various equilibrium allocations resulting from the fiscal shock. Compared to the stationary equilibrium, there is an average annual increase of 0.421% in employment and 1.598% in wages. Consequently, this leads to a $2.997 increase in earnings and $2.577 increase in consumption for every $1 of fiscal spending. In the partial equilibrium scenario without changes in factor prices, the increase in labor demand is significantly more pronounced. This results in a $3.382 surge in earnings.[27]

Table D.10: Short-run responses of other equilibrium allocations

	Baseline	Partial eq.
Employment (average annual % change from ss)	0.4210	0.6506
Wage (average annual % change from ss)	1.5983	
Earnings (dollars per $1 spending)	2.9970	3.3815
Comsumption (dollars per $1 spending)	2.5771	

D.6 Parameter-level uncertainty associated with the fiscal multipliers

we have examined the parameter-level uncertainty associated with the fiscal multipliers by drawing 50 draws from the estimated posterior distribution.[28] In the table below, the values in the square brackets represent the 10th and 90th percentiles of the fiscal multipliers, providing a sense of the range of possible outcomes.

We observe that the output fiscal multiplier in the heterogeneous firm model remains significantly above 1, whereas in the representative firm model, it falls

[27]The consumption is fixed at the steady-state level to turn off the inter-temporal bond price variation.

[28]Due to computational constraints, we limited the analysis to 50 draws.

below 1. Moreover, the investment response differs significantly between the two models. In the heterogeneous-firm model, investment exhibits a more muted decline than in the representative-firm counterpart, where the crowding-out effect is more pronounced. This suggests that firm heterogeneity plays a key role in shaping the transmission of fiscal policy, influencing both output and investment dynamics in ways that would not be captured in a representative-agent framework.

Table D.11: Posterior analysis of fiscal multipliers

	Heterogeneous-firm model	Representative-firm model
Output	1.1494	0.8906
	[1.0530, 1.6984]	[0.8840,0.9387]
Investment	-0.2055	-0.4371
	[-0.2082, -0.1924]	[-0.4455,-0.4370]

E Assumptions for the cross-state analysis

In this section, we specify the assumptions to implement the cross-state analysis using the baseline model. As there is only a representative household in this economy, the state-level consumption is defined under the following assumptions:

- All the incomes are state-specific, and there is no cross-state transfer.
- Each equity is exclusively owned by the state's household.
- Bond holding and lump-sum subsidies are attributed to each state proportionately to the exogenous fiscal spending ratio.

Given these assumptions, the state-level consumption can be properly defined due to the separate budget clearing across the states. One can introduce two households in the model to capture Poor and Good households separately, but this can be done only at a high computational cost and the model complication.

F Notes on the fiscal multiplier analysis

The following laws of motion determine the time path of the public capital stocks after the fiscal spending shock ΔG at $t = 1$:

$$\mathcal{N}_{A,t+1} = \mathcal{N}_{At}(1-\delta_{\mathcal{N}}) + \mathcal{F}_t - \frac{\mu}{2}\left(\frac{\mathcal{F}_t}{\mathcal{N}_{At}}\right)^2 \mathcal{N}_{At} \tag{31}$$

$$\mathcal{N}_{jt} = \zeta_j \mathcal{N}_{At} \quad \text{for } j \in \{G, P\} \tag{32}$$

$$\mathcal{F}_t = \begin{cases} \mathcal{F}^{ss} + \Delta\mathcal{G} & \text{if } t = 1 \\ \mathcal{F}^{ss} & \text{otherwise} \end{cases} \tag{33}$$

where $\mathcal{F}^{ss}$ is the stationary equilibrium level of infrastructure spending.

For a simple illustration, we consider a two-period model with the firm-level investment decision where the production functions are the same as in Proposition 1, and investment is subject to the convex adjustment cost. From the first-order condition of the investment, the following equation holds:

$$\underbrace{1 + \mu\left(\frac{k'}{k} - (1-\delta)\right)}_{\text{Marginal cost}} = \underbrace{\overbrace{\frac{1}{1+r}}^{\text{GE channel}} \overbrace{z^{\frac{1}{\alpha}}\left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}\left(\theta k'^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}-1} k'^{-\frac{1}{\lambda}}\theta}^{\text{Future MPK}}}_{\text{Marginal benefit = discounted future MPK}} \tag{34}$$

The left-hand side of the equation above is the marginal cost of the firm-level investment, and the right-hand side is the marginal benefit. To analyze how the increase in the public capital stock N affects the marginal benefit of firm-level investment, we a take a partial derivative with respect to N.

$$\frac{\partial}{\partial N}\text{Marginal benefit} = \left(\frac{1}{1+r}\right) \times \frac{\partial}{\partial N}\text{Future MPK} + \overbrace{\text{Future MPK} \times \frac{\partial}{\partial N}\left(\frac{1}{1+r}\right)}^{\text{GE effect}} \tag{35}$$

$$\frac{\partial}{\partial N}\text{Future MPK} = \frac{\partial}{\partial N}\left(\theta k'^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}}\right)^{\frac{1}{\lambda-1}} F(\Theta) \tag{36}$$

$$= \left(\frac{1}{\lambda-1}\right)\left(\frac{\lambda-1}{\lambda}\right)(1-\theta)N^{-\frac{1}{\lambda}}\left(\theta k'^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}}\right)^{\frac{2-\lambda}{\lambda-1}} F(\Theta) \tag{37}$$

$$= \frac{1}{\lambda}(1-\theta)N^{-\frac{1}{\lambda}}\left(\theta k'^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}}\right)^{\frac{2-\lambda}{\lambda-1}} F(\Theta) > 0 \tag{38}$$

where F is a function of the parameters, Θ. If the elasticity of substitution λ is a finite positive number, the marginal benefit of firm-level investment increases in N through the increased future marginal product of capital, given the general equilibrium effect is fixed. However, if λ goes to infinity, the marginal benefit of investment does not depend on N. It is worth noting that the marginal benefit increases in N regardless of whether the public and private capital stocks are gross complements ($\lambda < 1$) or substitutes ($\lambda > 1$).

The elasticity of substitution between private and public capital stock plays a key role in determining the marginal benefit of firm-level investment given a fiscal expenditure shock. Analytically, the change in the marginal benefit of firm-level investment over the elasticity given the fiscal spending shock can be captured by the cross derivative $\left[\frac{\partial^2}{\partial\lambda\partial N}\text{Marginal benefit}\right]$ in the simple two-period model.

Using Equation (35), we have the following equation:

$$\frac{\partial^2}{\partial\lambda\partial N}\text{Marginal benefit} = \frac{\partial}{\partial\lambda}\left[\underbrace{\frac{1}{\lambda}}_{\text{Direct}}\underbrace{\frac{\left(\frac{1}{1+r}\right)MPK}{\left(\theta k'\left(\frac{N}{k'}\right)^{\frac{1}{\lambda}}+(1-\theta)N\right)}}_{\text{Indirect}}\right]+\frac{\partial}{\partial\lambda}\text{GE effect} \tag{39}$$

As displayed in the equation above, the elasticity of substitution affects the response of marginal benefit through two channels: 1) direct and 2) indirect channels. The direct channel refers to newly added capital being relatively less valuable when the public capital stocks are more substitutable with the private capital. The indirect channel refers to a change in marginal benefit of investment due to the change in the relative values of the existing public and private capital stocks. The direct channel predicts the marginal benefit of firm-level investment decreases in the elasticity, while the sign of the indirect channel cannot be analytically determined.[29]

G The representative-agent model: An extension of Baxter and King (1993)

We consider the following representative-firm problem where the notations are the same as the baseline model except for ζ, which is the scale parameter for the infrastructure capital. It is worth noting that we use the same Φ level for the representative-agent model as in the baseline. This is to preserve the symmetry

[29]The sign of the effect also depends on the firm-level capital stock.

in the adjustment costs between the private sector and the public sector.[30] Also, the household and government sides are identical to the baseline model, so we abstract from the description for the sake of brevity.

$$J(k;S) = \max_{k'} \quad (1-\gamma)\left(\frac{\gamma}{w(S)}\right)^{\frac{\gamma}{1-\gamma}} A^{\frac{1}{1-\gamma}} N^{\frac{\zeta}{1-\gamma}} k^{\frac{\alpha}{1-\gamma}} (1-\tau^c)(1-\tau^h) \tag{40}$$

$$+ (-k' + (1-\delta)k)(1-\tau^h) + \tau\delta k(1-\tau^h) \tag{41}$$

$$- \frac{\mu}{2}\left(\frac{k'}{k} - (1-\delta)\right)^2 k(1-\tau^h) + \frac{1}{1+r(S)}\mathbb{E}J(k';S') \tag{42}$$

where J is the value of the representative firm; S is the aggregate state that include the same components as the baseline model's aggregate state, except for the distribution of capital Φ replaced by the aggregate capital stock K. The first-order optimality conditions are as follows:

$$[k'] : \quad \left(1 + \mu\left(\frac{k'}{k} - (1-\delta)\right)\right)(1-\tau^h) = \frac{1}{1+r(S)}\mathbb{E}J_1(k';S') \tag{43}$$

Also, from the envelope theorem, we have

$$[k] : \quad J_1(k;S) = \frac{\alpha}{1-\gamma}(1-\gamma)\left(\frac{\gamma}{w(S)}\right)^{\frac{\gamma}{1-\gamma}} A^{\frac{1}{1-\gamma}} N^{\frac{\zeta}{1-\gamma}} k^{\frac{\alpha}{1-\gamma}-1} (1-\tau^c)(1-\tau^h) \tag{44}$$

$$+ (1-\delta+\tau^c\delta)(1-\tau^h) + \left(\frac{\mu}{2}\left(\frac{k'}{k}\right)^2 - \frac{\mu}{2}(1-\delta)^2\right)(1-\tau^h). \tag{45}$$

[30]If the representative-agent economy's private capital adjustment cost is differently calibrated, it necessarily implies less or more efficient adjustment than the infrastructure capital adjustment. Also, it is not desirable to change the public capital adjustment cost parameter for the sake of a fair comparison of the fiscal multipliers across the models.

H Comparison with the fiscal multipliers in the literature

Ramey (2011) constructs government spending shocks that control for the anticipation effects and finds the government spending multipliers ranging from 0.6 to 1.2. Nakamura and Steinsson (2018) use state-level variations in military buildups (increases in federal purchases associated with military buildups) and find a state GDP multiplier of 1.4. They show their multiplier is equal to the aggregate multiplier in a small open economy with a fixed exchange rate and is larger than the closed economy aggregate multiplier for normal monetary policy. Chodorow-Reich (2019) studies the fiscal multiplier using the cross-sectional variations of fiscal spending controlling for state-specific heterogeneity. The paper concludes that the cross-sectional multiplier is around 1.8, and the lower bound for the national multiplier without the monetary policy response is 1.7. Our quantitative result shows that the short-run national fiscal multiplier is around 1.41 without the general equilibrium effect, comparable but smaller value than Chodorow-Reich (2019). However, once the general equilibrium effect (the real interest rate) is considered, the spending multiplier reduces to 1.15, which is within the range of Ramey (2011).

Our focus is on infrastructure spending multipliers, which may have a different nature compared to defense or overall government spending, as pointed out in Leduc and Wilson (2017). The paper empirically analyzes the fiscal multiplier using the state-level variation in highway spending and structurally analyzes the national-level multiplier based on the empirical analysis. They provide the impact multiplier of 1.4 and the cumulative multiplier of 6.6 over 10 years. Our paper distinguishes itself from Leduc and Wilson (2017) by including a firm-level frictional adjustment that accurately models micro-level investment responses after the implementation of fiscal policies (Cooper and Haltiwanger, 2006). We also estimate the micro-level parameters using the firm and state-level moments.

Fishback and Kachanovskaya (2015) discuss that the possible spillover effects across the states after fiscal spending hinder the translation of the cross-sectional multiplier analysis to the national-level multiplier. For this, recent literature has found a breakthrough by quantifying the spillover effects through the cross-state network (Peri, Rachedi, and Varotto, 2023). In our paper, we adopt a different approach to address this issue. We classify states into two groups — Good and Poor — based on geographical proximity, broadly distinguishing between the East and West. However, to mitigate potential spillover effects that could bias our analysis, we exclude neighboring states where infrastructure benefits may extend across regional boundaries. This grouping strategy ensures a representation of regional differences in infrastructure while minimizing cross-regional externalities.

I Robustness checks

I.1 Robustness over different congestion effects

We conduct the battery of robustness checks for the different levels of the congestion parameter ι. The parameter determines the eroded portion of output due to the congestion effect, and it affects the transition path of the Good and Poor regions in the post shock periods. The key channel for congestion effect is the firm-level incentives to invest — forward-looking firms have less incentive to invest as additional capital becomes less profitable.

We consider the different levels corresponding to the steady-state output erosion ranging from 1% to 10%. According to our fiscal multiplier calculations, the output fiscal multiplier strictly monotonically decreases in the congestion parameter. However the magnitude variation is negligibly small, marking the fiscal multiplier with the 10% steady-state output erosion by congestion at 1.1493, which is only 0.01% decrease from the baseline multiplier. The main reason we found for these small variations is the denominator effect — steady-state output is also strongly affected by congestion, and the levels across the whole post-shock transition are almost uniformly affected by congestion. Then, the shock response relative to the steady state level is not strongly affected by the congestion.

I.2 Robustness over a different exogenous spatial transition process

In this section, we discuss the robustness check over a different exogenous spatial transition process. In particular, we consider a counter-factual economy where the transition probability of a firm at the Good region increases by 1 percent (2% $\rightarrow$ 3%). In this economy, the portion of firms in the Good region is 0.7611 (0.8270 $\rightarrow$ 0.7611), which is significantly lower than the empirical estimate. Given this dramatic change, the impact on the fiscal multiplier is mild, leading to the level of 1.1018 (1.1494 $\rightarrow$ 1.1018).

In this exercise, by increasing the shuffling rate between the two areas, well-capitalized firms in the Good region move to the Poor region more often, leading to a positive impact in the Poor region and a negative impact on the Good region. Therefore, firms in the Good region have less incentive to invest, as there is a greater chance to move to the other area, while the firms in the Poor region have a greater incentive to invest. Given these two channels, there are countervailing forces for the fiscal multiplier, and it turns out that the increased shuffling probability slightly negatively affects the fiscal multiplier.

Consistent with these findings, when the transition probability of a firm at the Good region decreases by 1 percent ($2\% \rightarrow 1\%$), the fiscal multiplier slightly increases — the fiscal multiplier becomes 1.1550, compared to our baseline of 1.1494. The slight increase occurs because lower transition probabilities increase firms' investment incentives in the Good region due to reduced uncertainty about future productivity. These small changes confirm that our results are not driven by these transition assumptions.

I.3 Robustness over different Frisch elasticities

In this section, we compare output fiscal multipliers across different Frisch elasticities, holding all other parameters at their baseline calibrated values and posterior mean estimates. Lower Frisch elasticity leads to significantly lower fiscal multipliers: when the elasticity is 0.5, the fiscal multiplier falls to 0.731, while an elasticity of 5 yields a multiplier of 1.427. This monotonic relationship arises because lower Frisch elasticity (less elastic labor supply) creates stronger wage pressure to firms following positive demand shocks. When labor supply is relatively inelastic, increased labor demand from infrastructure spending drives up wages more sharply, raising firms' costs and dampening their investment response. This mechanism reduces the overall output multiplier, consistent with findings in Ramey (2020).

Table I.12: Fiscal multipliers over different Frisch elasticities

	$\chi = 0.5$	$\chi = 1$	$\chi = 3$	Baseline ($\chi = 4$)	$\chi = 5$
Output multiplier	0.731	0.813	0.980	1.149	1.427

J Proof of Propositions

Proposition 1. *Suppose we are given the micro-level data set* (k_1, k_2, y_1, y_2, N) *s.t.*

$$\exists i \in \{1,2\} \text{ s.t. } k_i < N, \quad N \leq k_1 + k_2, \quad \frac{y_1}{k_1} = \frac{y_2}{k_2}. \tag{46}$$

Suppose the micro-level estimates (z, λ) *and the aggregate-level estimate* ξ *are exactly identified by fitting the data with the production functions as follows:*

$$f(k_1, N; \lambda, z) = y_1 \tag{47}$$

$$f(k_2, N; \lambda, 1) = y_2. \tag{48}$$

$$f(k_1 + k_2, N; \xi, 1) = y_1 + y_2 \tag{49}$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$*, the aggregate-level input elasticity satisfies* $\xi < 1$.

Proof.

We prove the proposition separately for production functions of constant (CRS) and decreasing returns to scale (DRS). We start from the CRS case.[31]

CRS production function Without loss of generality suppose $k_1 > k_2$, $z > 1$, and let $k_2 < N$. From the production functions, we have

$$y_1 = z^{\frac{1}{\alpha}} B(\theta k_1^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{50}$$

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{51}$$

$$y_1 + y_2 = B(\theta (k_1 + k_2)^{\frac{\xi-1}{\xi}} + (1-\theta) N^{\frac{\xi-1}{\xi}})^{\frac{\xi}{\xi-1}} \tag{52}$$

[31] The proof for the DRS production function is general enough to encompass the CRS case, but the CRS proof is more intuitive. So, we include both proofs.

where $B := \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$. Therefore, the following relationships hold (from the second and the third equations above):

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{53}$$

$$\left(\frac{y_1+y_2}{B(k_1+k_2)}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}}. \tag{54}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\xi \geq 1$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{55}$$

Hence, $\frac{y_2}{Bk_2} > 1$. From the condition $\frac{y_1}{k_1} = \frac{y_2}{k_2}$,

$$1 < \frac{y_2}{Bk_2} = \frac{y_1+y_2}{B(k_1+k_2)}. \tag{56}$$

As $\xi \geq 1$, we have

$$1 < \left(\frac{y_1+y_2}{B(k_1+k_2)}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}}. \tag{57}$$

However, $N \leq k_1 + k_2$. Thus, $\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}} \leq 1$. This leads to

$$\theta + (1-\theta)\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}} \leq 1, \tag{58}$$

which is a contradiction. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$, then the aggregate-level input elasticity satisfies $\xi < 1$.

DRS production function Without loss of generality suppose $k_1 > k_2, z > 1$, and let $k_2 < N$. From the production functions, we have

$$y_1 = z^{\frac{1}{1-\gamma}} B(\theta k_1^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\frac{\alpha}{1-\gamma}} \tag{59}$$

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\frac{\alpha}{1-\gamma}} \tag{60}$$

$$y_1 + y_2 = B(\theta (k_1+k_2)^{\frac{\xi-1}{\xi}} + (1-\theta) N^{\frac{\xi-1}{\xi}})^{\frac{\xi}{\xi-1}\frac{\alpha}{1-\gamma}} \tag{61}$$

where $B := \left(\frac{\gamma}{w}\right)^{\frac{\gamma}{1-\gamma}}$. Therefore, the following relationships hold (from the second and the third equations above):

$$\left(\frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{62}$$

$$\left(\frac{(y_1+y_2)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} (k_1+k_2)}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}}. \tag{63}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\xi \geq 1$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{64}$$

Hence, $\frac{y_2^{\frac{1-\gamma}{\alpha}}}{Bk_2} > 1$. From $\frac{y_1}{k_1} = \frac{y_2}{k_2}$ and $k_1 > k_2$,

$$\frac{y_1^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_1} = \frac{\left(\frac{y_2}{k_2}k_1\right)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_1} = \frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2} k_2^{\frac{\alpha+\gamma-1}{\alpha}} k_1^{\frac{1-\alpha-\gamma}{\alpha}} = \frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2} \underbrace{\left(\frac{k_1}{k_2}\right)^{\frac{1-\alpha-\gamma}{\alpha}}}_{>1} > 1. \tag{65}$$

Thus, we deduce that $\frac{y_1^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}k_1} > 1$.

Then, we derive the following inequalities:[32]

$$\frac{(y_1+y_2)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}(k_1+k_2)} > \frac{y_1^{\frac{1-\gamma}{\alpha}}+y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}(k_1+k_2)} \geq min\left\{\frac{y_1^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}k_1}, \frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}k_2}\right\} > 1. \tag{67}$$

As $\xi \geq 1$, we have

$$1 < \left(\frac{(y_1+y_2)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}(k_1+k_2)}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}}. \tag{68}$$

However, $N \leq k_1 + k_2$. Thus, $\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}} \leq 1$. This leads to

$$\theta + (1-\theta)\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}} \leq 1, \tag{69}$$

which is a contradiction. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$, then the aggregate-level input elasticity satisfies $\xi < 1$. ■

Proposition 2. *Suppose we are given the micro-level data set (k_1, k_2, y_1, y_2, N) s.t.*

$$\exists i \in \{1,2\} \; s.t. \; k_i < N, \quad 1 < N \leq k_1 + k_2, \quad \frac{y_1}{k_1} = \frac{y_2}{k_2}. \tag{70}$$

Suppose the micro-level estimates $(\{z_i\}_{i=1}^n, \lambda)$ and the aggregate-level estimate ζ are ex-

[32]The first inequality is from the Jensen's inequality for concave function, as $\frac{1-\gamma}{\alpha} < 1$. The second inequality is from

$$\frac{a+b}{c+d} > min\left\{\frac{a}{c}, \frac{b}{d}\right\}, \tag{66}$$

which can be proven by the proof by contradiction.

actly identified by fitting the data with the production functions as follows:

$$f(k_1, N; \lambda, z) = y_1 \tag{71}$$

$$f(k_2, N; \lambda, 1) = y_2. \tag{72}$$

$$h(k_1 + k_2, N; \zeta, 1) = y_1 + y_2 \tag{73}$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$, the public capital scale parameter satisfies $\zeta > 0$.

Proof.

We prove the proposition separately for production functions of constant (CRS) and decreasing returns to scale (DRS). We start from the CRS case.[33]

CRS production function Without loss of generality suppose $k_1 > k_2, z > 1$, and let $k_2 < N$. From the production functions, we have

$$y_1 = z^{\frac{1}{\alpha}} B(\theta k_1^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{74}$$

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{75}$$

$$y_1 + y_2 = B(k_1 + k_2) N^{\frac{\zeta}{\alpha}} \tag{76}$$

where $B := \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$. Therefore, the following relationships hold (from the sec-

[33] The proof for the DRS production function is general enough to encompass the CRS case, but the CRS proof is more intuitive. So, we include both proofs.

ond and the third equations above):

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{77}$$

$$\frac{y_1+y_2}{B(k_1+k_2)} = N^{\frac{\zeta}{\alpha}}. \tag{78}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\zeta < 0$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{79}$$

Hence, $\frac{y_2}{Bk_2} > 1$. From the condition $\frac{y_1}{k_1} = \frac{y_2}{k_2}$,

$$1 < \frac{y_2}{Bk_2} = \frac{y_1+y_2}{B(k_1+k_2)}. \tag{80}$$

Thus, we have

$$1 < \frac{y_1+y_2}{B(k_1+k_2)} = N^{\frac{\zeta}{\alpha}}, \tag{81}$$

which is a contradiction, as $\zeta < 0$ and $N > 1$. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$ under the non-rivalry, then the public capital scale parameter satisfies $\zeta > 0$ (Baxter and King, 1993).

DRS production function Without loss of generality suppose $k_1 > k_2$, $z > 1$, and let $k_2 < N$. From the production functions, we have

$$y_1 = z^{\frac{1}{1-\gamma}} B(\theta k_1^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\frac{\alpha}{1-\gamma}} \tag{82}$$

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\frac{\alpha}{1-\gamma}} \tag{83}$$

$$y_1 + y_2 = B(k_1 + k_2)^{\frac{\alpha}{1-\gamma}} N^{\frac{\zeta}{1-\gamma}}, \tag{84}$$

where $B := \left(\frac{\gamma}{w}\right)^{\frac{\gamma}{1-\gamma}}$.

Therefore, the following relationships hold (from the second and the third equations above):

$$\left(\frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{85}$$

$$\frac{(y_1 + y_2)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}(k_1 + k_2)} = N^{\frac{\zeta}{\alpha}}. \tag{86}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\zeta < 0$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{87}$$

Hence, $\frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2} > 1$. From $\frac{y_1}{k_1} = \frac{y_2}{k_2}$ and $k_1 > k_2$,

$$\frac{y_1^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_1} = \frac{\left(\frac{y_2}{k_2} k_1\right)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_1} = \frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2} k_2^{\frac{\alpha+\gamma-1}{\alpha}} k_1^{\frac{1-\alpha-\gamma}{\alpha}} = \frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2} \underbrace{\left(\frac{k_1}{k_2}\right)^{\frac{1-\alpha-\gamma}{\alpha}}}_{>1} > 1. \tag{88}$$

Thus, we deduce that $\frac{y_1^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_1} > 1$.

Then, we derive the following inequalities:[34]

$$\frac{(y_1 + y_2)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}(k_1 + k_2)} > \frac{y_1^{\frac{1-\gamma}{\alpha}} + y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}(k_1 + k_2)} \geq min \left\{ \frac{y_1^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_1}, \frac{y_2^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}} k_2} \right\} > 1. \tag{90}$$

Thus, we have

$$1 < \frac{(y_1 + y_2)^{\frac{1-\gamma}{\alpha}}}{B^{\frac{1-\gamma}{\alpha}}(k_1 + k_2)} = N^{\frac{\zeta}{\alpha}}, \tag{91}$$

which is a contradiction, as $\zeta < 0$ and $N > 1$. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$ under the non-rivalry, then the public capital scale parameter satisfies $\zeta > 0$ (Baxter and King, 1993). ■

K A generalized simple theory

K.1 A simple theory with multiple (discrete) firms

Corollary 2. *Suppose we are given the micro-level data set* $\left(\{k_i, y_i\}_1^n, N\right)$ *s.t.*

$$\exists i \in \{1, 2, \ldots, n\} \text{ s.t. } k_i < N, \quad N \leq \sum_i k_i, \quad \frac{y_1}{k_1} = \frac{y_2}{k_2} = \cdots = \frac{y_n}{k_n}. \tag{92}$$

[34] The first inequality is from the Jensen's inequality for concave function, as $\frac{1-\gamma}{\alpha} < 1$. The second inequality is from

$$\frac{a+b}{c+d} > min\left\{\frac{a}{c}, \frac{b}{d}\right\}, \tag{89}$$

which can be proven by the proof by contradiction.

Suppose the micro-level estimates $(\{z_i\}_{i=1}^n, \lambda)$ *and the aggregate-level estimate* ξ *are exactly identified by fitting the data with the production functions as follows:*

$$f(k_i, N; \lambda, z_i) = y_i, \quad \textit{for } \forall i \in \{1, 2, \ldots, n\} \tag{93}$$

$$f\left(\sum_i k_i, N; \xi, 1\right) = \sum_i y_i \tag{94}$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$*, the aggregate-level input elasticity satisfies* $\xi < 1$.

Proof.

For brevity, we prove the corollary for the CRS production case. However, the extension to the DRS production function is identical to the one for Proposition 1.[35]

Without loss of generality suppose $k_i > k_2$ for $\forall i$, $z_2 = 1$ and let $k_2 < N$. From the production functions, we have

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{95}$$

$$\sum_i y_i = B\left(\theta\left(\sum_i k_i\right)^{\frac{\xi-1}{\xi}} + (1-\theta)N^{\frac{\xi-1}{\xi}}\right)^{\frac{\xi}{\xi-1}} \tag{96}$$

where $B := \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$. Therefore, the following relationships hold (from the sec-

[35]The proofs for the DRS production function is available upon request.

ond and the third equations above):

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{97}$$

$$\left(\frac{\sum_i y_i}{B\sum_i k_i}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{\sum_i k_i}\right)^{\frac{\xi-1}{\xi}}. \tag{98}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\xi \geq 1$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{99}$$

Hence, $\frac{y_2}{Bk_2} > 1$. From the condition $\frac{y_1}{k_1} = \frac{y_2}{k_2} = \cdots = \frac{y_n}{k_n}$,

$$1 < \frac{y_2}{Bk_2} = \frac{\sum_i y_i}{B\sum_i k_i}. \tag{100}$$

As $\xi \geq 1$, we have

$$1 < \left(\frac{\sum_i y_i}{B\sum_i k_i}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{\sum_i k_i}\right)^{\frac{\xi-1}{\xi}}. \tag{101}$$

However, $N \leq \sum_i k_i$. Thus, $\left(\frac{N}{\sum_i k_i}\right)^{\frac{\xi-1}{\xi}} \leq 1$. This leads to

$$\theta + (1-\theta)\left(\frac{N}{\sum_i k_i}\right)^{\frac{\xi-1}{\xi}} \leq 1, \tag{102}$$

which is a contradiction. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$, then the aggregate-level input elasticity satisfies $\xi < 1$. ■

Corollary 3. *Suppose we are given the micro-level data set* $\left(\{k_i, y_i\}_1^n, N\right)$ *s.t.*

$$\exists i \in \{1, 2, \ldots, n\} \text{ s.t. } k_i < N, \quad 1 < N \leq \sum_i k_i, \quad \frac{y_1}{k_1} = \frac{y_2}{k_2} = \cdots = \frac{y_n}{k_n}. \tag{103}$$

Suppose the micro-level estimates $(\{z_i\}_{i=1}^n, \lambda)$ *and the aggregate-level estimate* ξ *are exactly identified by fitting the data with the production functions as follows:*

$$f(k_i, N; \lambda, z_i) = y_i, \quad \text{for } \forall i \in \{1, 2, \ldots, n\} \tag{104}$$

$$f\left(\sum_i k_i, N; \xi, 1\right) = \sum_i y_i \tag{105}$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$*, the public capital scale parameter satisfies* $\zeta > 0$*.*

Proof.

For brevity, we prove the corollary for the CRS production case. However, the extension to the DRS production function is identical to the one for Proposition 2.[36]

Without loss of generality suppose $k_i > k_2$ for $\forall i$, $z_2 = 1$ and let $k_2 < N$. From the production functions, we have

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{106}$$

$$\sum_i y_i = B\left(\theta\left(\sum_i k_i\right)^{\frac{\xi-1}{\xi}} + (1-\theta)N^{\frac{\xi-1}{\xi}}\right)^{\frac{\xi}{\xi-1}} \tag{107}$$

where $B := \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$. Therefore, the following relationships hold (from the sec-

[36]The proofs for the DRS production function is available upon request.

ond and the third equations above):

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{108}$$

$$\frac{\sum_i y_i}{B\sum_i k_i} = N^{\frac{\zeta}{\alpha}}. \tag{109}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\zeta < 0$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{110}$$

Hence, $\frac{y_2}{Bk_2} > 1$. From the condition $\frac{y_1}{k_1} = \frac{y_2}{k_2} = \cdots = \frac{y_n}{k_n}$,

$$1 < \frac{y_2}{Bk_2} = \frac{\sum_i y_i}{B\sum_i k_i}. \tag{111}$$

Thus, we have

$$1 < \frac{\sum_i y_i}{B\sum_i k_i} = N^{\frac{\zeta}{\alpha}}, \tag{112}$$

which is a contradiction, as $\zeta < 0$ and $N > 1$. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$ under the non-rivalry, then the public capital scale parameter satisfies $\zeta > 0$ (Baxter and King, 1993). ■

K.2 A simple theory with the continuum of firms

Corollary 4. *Suppose we are given the micro-level data set (k_j, y_j, N), $j \in [0,1]$ s.t.*

$$\exists i \in [0,1] \ s.t. \ k_i < N, \quad N \leq \int_0^1 k_j dj, \quad \frac{y_j}{k_j} = C \in \mathbb{R}. \tag{113}$$

where C is a constant. Suppose the micro-level estimates (z_j, λ) and the aggregate-level estimate ξ are exactly identified by fitting the data with the production functions as follows:

$$(Normalizer) \quad z_0 = 1 \tag{114}$$

$$f(k_j, N; \lambda, z_j) = y_j \tag{115}$$

$$f\left(\int k_j dj, N; \xi, 1\right) = \int_0^1 y_j dj \tag{116}$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$, the aggregate-level input elasticity satisfies $\xi < 1$.

Proof.

For brevity, we prove the corollary for the CRS production case. However, the extension to the DRS production function is identical to the one for Proposition 1.[37]

Without loss of generality suppose $k_0 < N$. From the production functions, we have

$$y_0 = B(\theta k_0^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{117}$$

$$\int_0^1 y_j dj = B\left(\theta\left(\int k_j dj\right)^{\frac{\xi-1}{\xi}} + (1-\theta)N^{\frac{\xi-1}{\xi}}\right)^{\frac{\xi}{\xi-1}} \tag{118}$$

where $B := \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$. Therefore, the following relationships hold (from the sec-

[37] The proofs for the DRS production function is available upon request.

ond and the third equations above):

$$\left(\frac{y_0}{Bk_0}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_0}\right)^{\frac{\lambda-1}{\lambda}} \tag{119}$$

$$\left(\frac{\int_0^1 y_j dj}{B\left(\int_0^1 k_j dj\right)}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{\int_0^1 k_j dj}\right)^{\frac{\xi-1}{\xi}}. \tag{120}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\xi \geq 1$. As $N > k_0$, $\left(\frac{N}{k_0}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_0}{Bk_0}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_0}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{121}$$

Hence, $\frac{y_0}{Bk_0} > 1$. From the condition $\frac{y_j}{k_j} = C$,

$$1 < \frac{y_2}{Bk_2} = \frac{\int_0^1 y_j dj}{B\int_0^1 k_j dj}. \tag{122}$$

As $\xi \geq 1$, we have

$$1 < \left(\frac{\int_0^1 y_j dj}{B\left(\int_0^1 k_j dj\right)}\right)^{\frac{\xi-1}{\xi}} = \theta + (1-\theta)\left(\frac{N}{\int k_j dj}\right)^{\frac{\xi-1}{\xi}}. \tag{123}$$

However, $N \leq \int k_j dj$. Thus, $\left(\frac{N}{\int k_j dj}\right)^{\frac{\xi-1}{\xi}} \leq 1$. This leads to

$$\theta + (1-\theta)\left(\frac{N}{\int k_j dj}\right)^{\frac{\xi-1}{\xi}} \leq 1, \tag{124}$$

which is a contradiction. Therefore, if the micro-level input elasticity satisfies $\lambda \geq$

1, then the aggregate-level input elasticity satisfies $\xi < 1$. ■

Corollary 5. *Suppose we are given the micro-level data set (k_j, y_j, N), $j \in [0,1]$ s.t.*

$$\exists i \in [0,1] \; s.t. \; k_i < N, \quad 1 < N \leq \int_0^1 k_j dj, \quad \frac{y_j}{k_j} = C \in \mathbb{R}. \tag{125}$$

where C is a constant. Suppose the micro-level estimates (z_j, λ) and the aggregate-level estimate η are exactly identified by fitting the data with the production functions as follows:

$$(Normalizer) \quad z_0 = 1 \tag{126}$$

$$f(k_j, N; \lambda, z_j) = y_j \tag{127}$$

$$h\left(\int k_j dj, N; \eta, 1\right) = \int_0^1 y_j dj \tag{128}$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$, the public capital scale parameter satisfies $\eta > 0$.

Proof.

For brevity, we prove the corollary for the CRS production case. However, the extension to the DRS production function is identical to the one for Proposition 2.[38]

Without loss of generality suppose $k_0 < N$. From the production functions, we have

$$y_0 = B(\theta k_0^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{129}$$

$$\int y_j dj = B\left(\int k_j dj\right) N^{\frac{\eta}{\alpha}} \tag{130}$$

where $B := \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$. Therefore, the following relationships hold (from the sec-

[38]The proofs for the DRS production function is available upon request.

ond and the third equations above):

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{131}$$

$$\frac{\int y_j dj}{B\int k_j dj} = N^{\frac{\eta}{\alpha}}. \tag{132}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\eta < 0$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{133}$$

Hence, $\frac{y_2}{Bk_2} > 1$. From the condition $\frac{y_1}{k_1} = \frac{y_2}{k_2}$,

$$1 < \frac{y_2}{Bk_2} = \frac{\int y_j dj}{B\int k_j dj}. \tag{134}$$

Thus, we have

$$1 < \frac{\int y_j dj}{B\int k_j dj} = N^{\frac{\eta}{\alpha}}, \tag{135}$$

which is a contradiction, as $\eta < 0$ and $N > 1$. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$ under the non-rivalry, then the public capital scale parameter satisfies $\eta > 0$ (Baxter and King, 1993). ■

K.3 A simple theory with congestion effect

Consider a CES production function $F(K, N, L; \lambda, z)$ with constant returns to scale (CRS):

$$F(k, N, l; \lambda, z, Y) = zG(Y)(\theta k^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\alpha} l^{1-\alpha}. \tag{136}$$

where $G(Y) = (\overline{Y}/Y)^{\psi}$ captures the congestion effect. Then the production function with the implicit labor demand is as follows:

$$f(k, N; \lambda, z, Y) = z^{\frac{1}{\alpha}} (\overline{Y}/Y)^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} (\theta k^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\alpha}. \tag{137}$$

Corollary 6. *Under the congestion effect represented by $G = G(Y)$ and constant returns to scale production function f, the effective aggregate production function $h = h(K, N; \lambda)$ that satisfies the following equation features decreasing returns to scale with respect to K.*

$$Y = f(K, N; \lambda, 1, Y) = h(K, N; \lambda) \tag{138}$$

Proof.

$$\begin{aligned} Y &= f(K, N; \lambda, 1, Y) && (139)\\ &= (\overline{Y}/Y)^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} (\theta K^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\alpha}. && (140) \end{aligned}$$

Then, by rearranging the terms with respect to Y,

$$
\begin{aligned}
Y &= h(K, N; \lambda) && (141)\\
&= \left(\overline{Y}^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}} (\theta K^{\frac{\lambda-1}{\lambda}} + (1-\theta)N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}\alpha}\right)^{\frac{\alpha}{\alpha+\psi}}. && (142)
\end{aligned}
$$

As $\frac{\alpha}{\alpha+\psi} \in (0,1)$, h is a decreasing-returns-to-scale production function. ■

Corollary 7. *Suppose we are given the micro-level data set (k_1, k_2, y_1, y_2, N) s.t.*

$$\exists i \in \{1,2\} \text{ s.t. } k_i < N, \quad 1 < N \leq k_1 + k_2, \quad \frac{y_1}{k_1} = \frac{y_2}{k_2}. \tag{143}$$

Suppose the micro-level estimates (z, λ) and the aggregate-level estimate ξ are exactly identified by fitting the data with the production functions as follows:

$$
\begin{aligned}
f(k_1, N; \lambda, z, Y) &= y_1 && (144)\\
f(k_2, N; \lambda, 1, Y) &= y_2. && (145)\\
h(k_1 + k_2, N; \xi, 1) &= Y = y_1 + y_2 && (146)
\end{aligned}
$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$, the aggregate-level input elasticity satisfies $\xi < 1$.

Proof.

For brevity, we prove the corollary for the CRS production case. However, the extension to the DRS production function is identical to the one for Proposition 1.[39]

Without loss of generality suppose $k_1 > k_2$, $z > 1$, and let $k_2 < N$. From the

[39]The proofs for the DRS production function is available upon request.

production functions, we have

$$y_1 = z^{\frac{1}{\alpha}} B(\theta k_1^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{147}$$

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{148}$$

$$y_1 + y_2 = \left(\widetilde{B}(\theta (k_1+k_2)^{\frac{\xi-1}{\xi}} + (1-\theta) N^{\frac{\xi-1}{\xi}})^{\frac{\xi}{\xi-1}} \right)^{\frac{\alpha}{\alpha+\psi}} \tag{149}$$

where $B := (\overline{Y}/Y)^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$ and $\widetilde{B} := \overline{Y}^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$. Therefore, the following relationships hold (from the second and the third equations above):

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{150}$$

$$\left(\frac{y_1+y_2}{\widetilde{B}^{\frac{\alpha}{\alpha+\psi}}(k_1+k_2)^{\frac{\alpha}{\alpha+\psi}}}\right)^{\frac{\xi-1}{\xi}\frac{\alpha+\psi}{\alpha}} = \theta + (1-\theta)\left(\frac{N}{k_1+k_2}\right)^{\frac{\xi-1}{\xi}}. \tag{151}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\xi \geq 1$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{152}$$

Hence, $\frac{y_2}{Bk_2} > 1$.

Then, from the condition $\frac{y_1}{k_1} = \frac{y_2}{k_2}$,

$$1 < \frac{y_2}{Bk_2} = \frac{y_1+y_2}{B(k_1+k_2)} = \frac{Y^{\frac{\alpha+\psi}{\alpha}}}{\widetilde{B}(k_1+k_2)} = \left(\frac{y_1+y_2}{\widetilde{B}^{\frac{\alpha}{\alpha+\psi}}(k_1+k_2)^{\frac{\alpha}{\alpha+\psi}}}\right)^{\frac{\alpha+\psi}{\alpha}}. \tag{153}$$

As $\xi \geq 1$, we have

$$1 < \left(\frac{y_1 + y_2}{\widetilde{B}^{\frac{\alpha}{\alpha+\psi}} (k_1 + k_2)^{\frac{\alpha}{\alpha+\psi}}} \right)^{\frac{\xi-1}{\xi} \frac{\alpha+\psi}{\alpha}} = \theta + (1-\theta) \left(\frac{N}{k_1 + k_2} \right)^{\frac{\xi-1}{\xi}}. \tag{154}$$

However, $N \leq k_1 + k_2$. Thus, $\left(\frac{N}{k_1+k_2} \right)^{\frac{\xi-1}{\xi}} \leq 1$. This leads to

$$\theta + (1-\theta) \left(\frac{N}{k_1 + k_2} \right)^{\frac{\xi-1}{\xi}} \leq 1, \tag{155}$$

which is a contradiction. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$, then the aggregate-level input elasticity satisfies $\xi < 1$. ■

Then, we show that Proposition 2 holds under a production function with a congestion effect. Consider a Cobb-Douglas production function $F(K, N, L; \lambda, z)$ with constant returns to scale (CRS):

$$H(k, N, L; \zeta, z, Y) = zG(Y)k^{\alpha} L^{1-\alpha} N^{\zeta} \tag{156}$$

where $G(Y) = (\overline{Y}/Y)^{\psi}$ captures the congestion effect. Then the production function with the implicit labor demand is as follows:

$$h(k, N; \zeta, z, Y) = z^{\frac{1}{\alpha}} \left(\frac{\overline{Y}}{Y} \right)^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w} \right)^{\frac{1-\alpha}{\alpha}} kN^{\frac{\zeta}{\alpha}} \tag{157}$$

The aggregate production is as follows:

$$Y = z^{\frac{1}{\alpha}} \left(\frac{\overline{Y}}{Y} \right)^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w} \right)^{\frac{1-\alpha}{\alpha}} KN^{\frac{\zeta}{\alpha}} \tag{158}$$

Then, the effective aggregate production after the congestion is as follows:

$$Y = g(K, N; \zeta, z) \tag{159}$$

$$= \left(z^{\frac{1}{\alpha}} \overline{Y}^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w} \right)^{\frac{1-\alpha}{\alpha}} K N^{\frac{\zeta}{\alpha}} \right)^{\frac{\alpha}{\alpha+\psi}} \tag{160}$$

which is a decreasing-returns-to-scale production function as predicted by Corollary 1. The following proposition shows that the original proposition holds even after accounting for the congestion effect.

Corollary 8. *Suppose we are given the micro-level data set* (k_1, k_2, y_1, y_2, N) *s.t.*

$$\exists i \in \{1, 2\} \text{ s.t. } k_i < N, \quad 1 < N \leq k_1 + k_2, \quad \frac{y_1}{k_1} = \frac{y_2}{k_2}. \tag{161}$$

Suppose the micro-level estimates (z, λ) *and the aggregate-level estimate* ζ *are exactly identified by fitting the data with the production functions as follows:*

$$f(k_1, N; \lambda, z, Y) = y_1 \tag{162}$$

$$f(k_2, N; \lambda, 1, Y) = y_2. \tag{163}$$

$$g(k_1 + k_2, N; \zeta, 1) = y_1 + y_2 \tag{164}$$

Then, if the micro-level input elasticity satisfies $\lambda \geq 1$*, the public capital scale parameter satisfies* $\zeta > 0$*.*

Proof.

For brevity, we prove the corollary for the CRS production case. However, the extension to the DRS production function is identical to the one for Proposition 2.[40]

[40] The proofs for the DRS production function is available upon request.

Without loss of generality suppose $k_1 > k_2$, $z > 1$, and let $k_2 < N$. From the production functions, we have

$$y_1 = z^{\frac{1}{\alpha}} B(\theta k_1^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{165}$$

$$y_2 = B(\theta k_2^{\frac{\lambda-1}{\lambda}} + (1-\theta) N^{\frac{\lambda-1}{\lambda}})^{\frac{\lambda}{\lambda-1}} \tag{166}$$

$$y_1 + y_2 = \left(\widetilde{B}(k_1 + k_2) N^{\frac{\zeta}{\alpha}}\right)^{\frac{\alpha}{\alpha+\psi}} \tag{167}$$

where $B := (\overline{Y}/Y)^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$ and $\widetilde{B} := \overline{Y}^{\frac{\psi}{\alpha}} \left(\frac{1-\alpha}{w}\right)^{\frac{1-\alpha}{\alpha}}$.

Therefore, the following relationships hold (from the second and the third equations above):

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} \tag{168}$$

$$\frac{(y_1 + y_2)^{\frac{\alpha+\psi}{\alpha}}}{\widetilde{B}(k_1 + k_2)} = N^{\frac{\zeta}{\alpha}}. \tag{169}$$

Suppose we are given $\lambda \geq 1$. We will prove the proposition by contradiction, so we assume $\zeta < 0$. As $N > k_2$, $\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1$. Thus,

$$\left(\frac{y_2}{Bk_2}\right)^{\frac{\lambda-1}{\lambda}} = \theta + (1-\theta)\left(\frac{N}{k_2}\right)^{\frac{\lambda-1}{\lambda}} > 1. \tag{170}$$

Hence, $\frac{y_2}{Bk_2} > 1$. From the condition $\frac{y_1}{k_1} = \frac{y_2}{k_2}$,

$$1 < \frac{y_2}{Bk_2} = \frac{y_1 + y_2}{B(k_1 + k_2)} = \frac{(y_1 + y_2)^{\frac{\alpha+\psi}{\alpha}}}{\widetilde{B}(k_1 + k_2)}. \tag{171}$$

Thus, we have

$$1 < \frac{(y_1 + y_2)^{\frac{\alpha+\psi}{\alpha}}}{\widetilde{B}(k_1 + k_2)} = N^{\frac{\zeta}{\alpha}}, \tag{172}$$

which is a contradiction, as $\zeta < 0$ and $N > 1$. Therefore, if the micro-level input elasticity satisfies $\lambda \geq 1$ under the non-rivalry, then the public capital scale parameter satisfies $\zeta > 0$ (Baxter and King, 1993). ■

K.4 Relaxing the assumption on the firm-level output-to-capital ratio

The proofs of propositions takes the following step, to which the assumption of $y_1/k_1 = y_2/k_2$ is applied:

$$1 < \frac{y_2}{Bk_2} = \frac{y_1 + y_2}{B(k_1 + k_2)}. \tag{173}$$

However, as long as $\frac{y_1+y_2}{B(k_1+k_2)} > 1$ is guaranteed, $\frac{y_1}{k_1} = \frac{y_2}{k_2}$ is not necessary. In this section, we drive the boundary of y_1/k_1 that assures the desired inequality.

Define $m := \frac{y_1}{k_1}$. Suppose $\frac{y_2}{Bk_2} = 1 + \chi$ where $\chi > 0$. We want to show that there exists $\overline{m} > 0$ such that if $m = \overline{m}$, $\frac{y_1+y_2}{B(k_1+k_2)} = 1$. Then, as long as $\frac{y_1}{k_1} = m > \overline{m}$, the desired inequality holds: $\frac{y_1+y_2}{B(k_1+k_2)} > 1$ even if $\frac{y_1}{k_1} \neq \frac{y_2}{k_2}$.

Using $y_1 = mk_1$ and $y_2 = Bk_2(1 + \chi)$,

$$\left.\frac{y_1 + y_2}{B(k_1 + k_2)}\right|_{m=\overline{m}} = \frac{\overline{m}k_1 + Bk_2(1 + \chi)}{B(k_1 + k_2)} = 1 \tag{174}$$

Rearranging the terms, we obtain

$$\overline{m} = B\left(1 - \chi\frac{k_2}{k_1}\right).$$

As long as $\frac{y_1}{k_1} > B\left(1 - \chi\frac{k_2}{k_1}\right)$, the desired inequality holds. It is worth noting that when $\frac{y_2}{Bk_2}$ is further away from 1 (χ is greater), the $\overline{m}$ becomes smaller, meaning the boundary becomes more lenient. In our setup, $\overline{m} = B\left(1 - \chi\frac{k_2}{k_1}\right) < 1$ for any choices of $\frac{k_2}{k_1} > 0$, as $B < 1$. Specifically,

$$\overline{m} \in [0, B), \quad \textit{where } \mathrm{B} \approxeq 0.15.$$

In our calibrated baseline model, $B \approxeq 0.15$. Suppose $\frac{y_2}{k_2} \approxeq 0.47$, which is from a standard neoclassical model's level ($\frac{Y}{K} = \frac{r+\delta}{\alpha}$). From this, we obtain $\chi \approxeq 2.1$. Then, regardless of $\frac{k_2}{k_1}$, as long as $\frac{y_1}{k_1} > 0.15 \approxeq B$, the desired inequality holds.

For a more concrete example, suppose $k_2 = 1.00$, $k_1 = 6.00$, and $y_2 = 2.80$, thus $\frac{y_2}{k_2} = 0.47$. Then, $\chi \approx 2.1$. Therefore, $\overline{m} \approx 0.10$. As long as $\frac{y_1}{k_1} > 0.10$, $\frac{y_1+y_2}{B(k_1+k_2)} > 1$ holds.

L Grouping relevance for fiscal multipliers

L.1 Theoretical results on grouping and aggregation bias

This section formalizes why grouping matters when productivity and infrastructure spending are positively sorted across states.

To provide sharp theoretical insights, we introduce a simple setup that shares the core ingredients with our baseline quantitative model. We assume N states, each featuring different productivity levels and infrastructure spending. Consider the following simple output function where A_j is the productivity and F_j is the infrastructure spending for state $j \in S = \{1, 2, \ldots, N\}$.

$$y_j = f(A_j, F_j) := A_j F_j, \tag{175}$$

where output is simply the product of productivity and infrastructure spending. In this setup, there is no distinction between the output level and the response of the output, as the model is static. Therefore, the output multiplier simply refers to the output levels.

Then, we consider a partition $\{G, P\}$ of S such that i) $S \setminus G = P$, ii) $|G| = |P|$, where $|X|$, $X \subseteq S$ denotes the number of states (elements) in set group X, and iii) for $\forall j \in G$ and $k \in P$, $F_j > F_k$ — Good and Poor groups are divided by the infrastructure spending size.[41] Thus, the partition $\{G, P\}$ is exactly the same as the grouping strategy as in our main text. The group-level output Y_G and Y_P are defined as follows:

$$Y_G := A_G F_G = \frac{\sum_{j \in G} A_j}{|G|} \frac{\sum_{j \in G} F_j}{|G|}, \quad Y_P := A_P F_P = \frac{\sum_{j \in P} A_j}{|P|} \frac{\sum_{j \in P} F_j}{|P|}, \tag{176}$$

It is worth noting that the production inputs are a group-level average productivity and infrastructure spending. The region-specific productivity in our baseline model is identical to this group-level average productivity.

We call the *positive sorting* between productivity and the infrastructure spending at the state level, when the following condition holds:

$$A_G > A_P. \tag{177}$$

[41] We assume N is an even number.

In other words, when the group-level average productivity is higher for the high infrastructure spending group, there is a positive sorting. There are other possible sorting measures, but this definition is simple and adequately serves the purpose of this section. According to this definition, our baseline model's estimates indicate positive sorting.

Specifically, we prove the following:[42]

- **Proposition 3 (Grouping relevance I).**

  If $A_G \geq A_P$, then

$$\frac{\sum_{j \in S} A_j}{N} \frac{\sum_{j \in S} F_j}{N} \leq (Y_G + Y_P)/2 \leq \frac{\sum_{j \in S} A_j F_j}{N} \tag{178}$$

  for the partition $\{G, P\}$ of S.

The implication of Proposition 3 is that state-by-state aggregation yields the maximal fiscal multiplier, while single-group aggregation produces the minimal fiscal multiplier. Our two-group aggregation necessarily falls between these bounds. Given this result, our primary objective was to identify the grouping strategy {G,P} that would yield the fiscal multiplier closest to the state-by-state maximum—that is, the least biased estimate of the true decentralized multiplier.

To sharpen our theoretical point, we now consider a stronger sorting condition to further characterize the optimality of our grouping strategy:

$$F_j > F_k \implies A_j > A_k. \tag{179}$$

If the condition above is satisfied, we call the *strict positive sorting* is given between A_j and F_j. The following proposition shows that the partition $\{G, P\}$ achieves the

[42]Each proposition's proof is provided in Appendix L.3.1 below.

highest level of output (multiplier) when there is strict positive sorting - therefore, it achieves the closest output multiplier to the aggregation of the state-level output multipliers.

- **Proposition 4 (Grouping relevance II).**

  Under the strict positive sorting between A_j and F_j,

$$(Y_M + Y_{M^c})/2 \leq (Y_G + Y_P)/2 \tag{180}$$

  for any evenly split partition $\{M, M^c\}$ of S.

The key distinction between the group-level and the state-by-state aggregation is that group-level aggregation cannot sharply capture the sorting between the spending and the productivity that happens at the state level — $\mathbb{E}(FA) = \mathbb{E}F\mathbb{E}A + cov(F, A)$, and group-level aggregation $\mathbb{E}F\mathbb{E}A$ misses $cov(F, A)$.

L.2 Empirical sorting pattern from the state-level data

In this section, we document that state-level productivity and infrastructure spending exhibit a clear positive sorting pattern over our sample period (1994–2017). Motivated by our theoretical results, it is natural to construct a regional grouping scheme that reflects this sorting pattern.

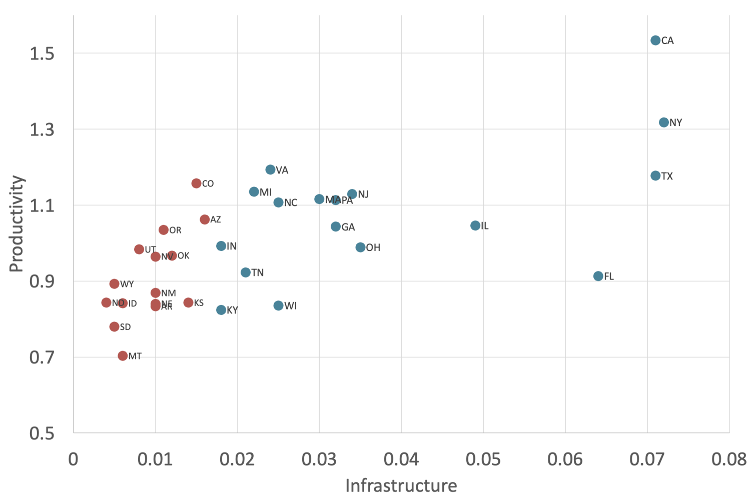
We estimate the parameters of state-level production functions following An, Kangur, and Papageorgiou (2019). Using nonlinear least squares, we fit the production function for state j in period t:

$$Y_{jt} = L_{jt}^{1-a}\left(bK_{jt}^{\psi} + (1-b)N_{jt}^{\psi}\right)^{\frac{a}{\psi}},$$

and obtain the estimates $a = 0.402$, $b = 0.07$, and $\frac{1}{1-\psi} = 0.445$. We define state-level productivity as the ratio of observed output to the fitted output from the estimated production function in each period, and then average across years to obtain the measure used in the figure (standardized to have a mean of one).

Figure L.3 plots this productivity measure against each state's share of total infrastructure capital stock. The positive correlation is evident (0.687): states with higher productivity tend to be associated with a larger share of infrastructure capital. The color coding reflects our regional classification, with "Poor" regions in brown and "Good" regions in green. This grouping aligns well with the observed sorting pattern. For instance, Texas and California are classified as "Good" regions because their high productivity coincides with substantial infrastructure spending, despite their western geographic location.

Figure L.3: State-level productivity versus infrastructure capital share



To summarize, our theoretical analysis formally proves that asymmetric grouping under positive sorting produces the most accurate (least biased) aggregation of fiscal multipliers. And our grouping is data-driven by positive sorting between productivity and infrastructure, not designed to maximize multipliers.

L.3 Proofs for propositions and a simple example

L.3.1 Proofs for propositions

Proposition 3 (Grouping relevance I)**.**

If $A_G \geq A_P$, then

$$\frac{\sum_{j\in S} A_j}{N}\frac{\sum_{j\in S} F_j}{N} \leq (Y_G + Y_P)/2 \leq \frac{\sum_{j\in S} A_j F_j}{N} \tag{181}$$

for the partition $\{G, P\}$ of S.

Proof.

We prove each inequality separately.

Part 1: Right Inequality. We need to show that $\frac{Y_G+Y_P}{2} \leq \frac{\sum_{j\in S} A_j F_j}{N}$.

By the Cauchy-Schwarz inequality, for any finite sets of real numbers $\{a_i\}$ and $\{b_i\}$:

$$\left(\sum_i a_i\right)\left(\sum_i b_i\right) \leq n \sum_i a_i b_i$$

where n is the number of elements.

Applying this to each group,

$$\left(\sum_{j\in G} A_j\right)\left(\sum_{j\in G} F_j\right) \leq |G| \sum_{j\in G} A_j F_j \tag{182}$$

$$\left(\sum_{j\in P} A_j\right)\left(\sum_{j\in P} F_j\right) \leq |P| \sum_{j\in P} A_j F_j \tag{183}$$

Therefore,

$$Y_G = \frac{\left(\sum_{j\in G} A_j\right)\left(\sum_{j\in G} F_j\right)}{|G|^2} \leq \frac{\sum_{j\in G} A_j F_j}{|G|} \tag{184}$$

$$Y_P = \frac{\left(\sum_{j\in P} A_j\right)\left(\sum_{j\in P} F_j\right)}{|P|^2} \leq \frac{\sum_{j\in P} A_j F_j}{|P|} \tag{185}$$

Thus,

$$\frac{Y_G + Y_P}{2} \leq \frac{1}{2}\left(\frac{\sum_{j\in G} A_j F_j}{|G|} + \frac{\sum_{j\in P} A_j F_j}{|P|}\right) \tag{186}$$

Then,

$$\frac{1}{2}\left(\frac{\sum_{j\in G} A_j F_j}{|G|} + \frac{\sum_{j\in P} A_j F_j}{|P|}\right) = \frac{\sum_{j\in G} A_j F_j}{N} + \frac{\sum_{j\in P} A_j F_j}{N} \tag{187}$$

$$= \frac{\sum_{j\in S} A_j F_j}{N} \tag{188}$$

Therefore:

$$\frac{Y_G + Y_P}{2} \leq \frac{1}{2}\left(\frac{\sum_{j\in G} A_j F_j}{|G|} + \frac{\sum_{j\in P} A_j F_j}{|P|}\right) = \frac{\sum_{j\in S} A_j F_j}{N}$$

Part 2: Left Inequality. We need to show that $\frac{\sum_{j\in S} A_j}{N}\frac{\sum_{j\in S} F_j}{N} \leq \frac{Y_G+Y_P}{2}$.

Since $|G| = |P| = \frac{N}{2}$, recall that

$$A_G = \frac{\sum_{j \in G} A_j}{|G|} = \frac{2}{N} \sum_{j \in G} A_j \quad \text{(average } A \text{ in group } G\text{)} \tag{189}$$

$$A_P = \frac{\sum_{j \in P} A_j}{|P|} = \frac{2}{N} \sum_{j \in P} A_j \quad \text{(average } A \text{ in group } P\text{)} \tag{190}$$

$$F_G = \frac{\sum_{j \in G} F_j}{|G|} = \frac{2}{N} \sum_{j \in G} F_j \quad \text{(average } F \text{ in group } G\text{)} \tag{191}$$

$$F_P = \frac{\sum_{j \in P} F_j}{|P|} = \frac{2}{N} \sum_{j \in P} F_j \quad \text{(average } F \text{ in group } P\text{)} \tag{192}$$

Then,

$$\frac{\sum_{j \in S} A_j}{N} = \frac{\sum_{j \in G} A_j + \sum_{j \in P} A_j}{N} = \frac{|G| \cdot A_G + |P| \cdot A_P}{N} = \frac{A_G + A_P}{2} \tag{193}$$

$$\frac{\sum_{j \in S} F_j}{N} = \frac{\sum_{j \in G} F_j + \sum_{j \in P} F_j}{N} = \frac{|G| \cdot F_G + |P| \cdot F_P}{N} = \frac{F_G + F_P}{2} \tag{194}$$

$$\frac{Y_G + Y_P}{2} = \frac{A_G F_G + A_P F_P}{2} \tag{195}$$

We need to prove:

$$\frac{A_G + A_P}{2} \cdot \frac{F_G + F_P}{2} \leq \frac{A_G F_G + A_P F_P}{2} \tag{196}$$

Multiplying both sides by 4:

$$(A_G + A_P)(F_G + F_P) \leq 2(A_G F_G + A_P F_P) \tag{197}$$

Expanding the left side:

$$A_G F_G + A_G F_P + A_P F_G + A_P F_P \leq 2 A_G F_G + 2 A_P F_P \tag{198}$$

Rearranging:

$$A_G F_P + A_P F_G \leq A_G F_G + A_P F_P \tag{199}$$

This is equivalent to:

$$0 \leq A_G F_G - A_G F_P - A_P F_G + A_P F_P = (A_G - A_P)(F_G - F_P) \tag{200}$$

By construction, $F_G > F_P$, and the positive sorting provides $A_G \geq A_P$. Therefore, the left-hand side inequality is proved.

■

The following corollary shows how the sorting affects this aggregation outcome by checking the inequality under $A_G = A_P$ (symmetric productivity).

Corollary 9 (Aggregation neutrality under the equal productivity)**.**
If $A_G = A_P$, then

$$\frac{\sum_{j\in S} A_j}{N} \frac{\sum_{j\in S} F_j}{N} = \frac{Y_G + Y_P}{2} \tag{201}$$

for the partition $\{G, P\}$ of S.

Proof.
If $A_G = A_P$, from Line (200) in the proof of Proposition 3,

$$(A_G - A_P)(F_G - F_P) = 0 \implies \frac{\sum_{j\in S} A_j}{N} \frac{\sum_{j\in S} F_j}{N} = \frac{Y_G + Y_P}{2} \tag{202}$$

Thus,

$$(Y_G + Y_P)/2 = \frac{\sum_{j \in S} A_j F_j}{N}. \tag{203}$$

■

So, when there is no group-level sorting ($A_G = A_P$), the aggregations of the representative and the group-level outputs are the same.

We now consider a stronger sorting condition to further characterize the optimality of our grouping strategy:

$$F_j > F_k \implies A_j > A_k. \tag{204}$$

If the condition above is satisfied, we call the *strict positive sorting* is given between A_j and F_j. The following proposition shows that the partition $\{G, P\}$ achieves the highest level of output (multiplier) when there is strict positive sorting - therefore, it achieves the closest output multiplier to the aggregation of the state-level output multipliers.

Proposition 4 (Grouping relevance II)**.**
Under the strict positive sorting between A_j and F_j,

$$(Y_M + Y_{M^c})/2 \leq (Y_G + Y_P)/2 \tag{205}$$

for any evenly split partition $\{M, M^c\}$ of S.

Proof.
Without loss of generality, we index F_j to satisfy $F_1 < F_2 < \cdots < F_N$. The strict positive sorting leads to $A_1 < A_2 < \cdots < A_N$. The partition $\{G, P\}$ guarantees

that

$$
\begin{aligned}
& F_G \geq F_M \geq F_P \\
& A_G \geq A_M \geq A_P
\end{aligned} \tag{206}
$$

for any evenly split partition $\{M, M^c\}$ of S.

Define $\overline{F} := \frac{F_G+F_P}{2} = \frac{F_M+F_{M^c}}{2}$ and $\overline{A} := \frac{A_G+A_P}{2} = \frac{A_M+A_{M^c}}{2}$. Then, for any given partition $\{M, M^c\}$, $(F_M, F_{M^c}, A_M, A_{M^c})$ could be written as

$$F_M = \overline{F} + x_M, \quad F_{M^c} = \overline{F} - x_M \tag{207}$$

$$A_M = \overline{A} + y_M, \quad A_{M^c} = \overline{A} - y_M \tag{208}$$

for some real numbers x_M and y_M (spreads).

Then,

$$A_M F_M + A_{M^c} F_{M^c} = (\overline{F} + x_M)(\overline{A} + y_M) + (\overline{F} - x_M)(\overline{A} - y_M) \tag{209}$$

$$= 2\overline{F}\overline{A} + 2x_M y_M \tag{210}$$

which strictly increases in x_M and y_M.

According to Line (206), the partition $\{G, P\}$ is the one with the largest spreads from the average levels: $x_G > x_M$ and $y_G > y_M$ for any partition $\{M, M^c\}$. Thus, we have

$$Y_M + Y_{M^c} = A_M F_M + A_{M^c} F_{M^c} \leq A_G F_G + A_P F_P = Y_G + Y_P \tag{211}$$

■

Therefore, our baseline grouping strategy yields the output multiplier closest to the first-best—namely, the aggregation of decentralized state-level multipliers. This is why we adopt this specific grouping strategy, as alternative approaches including symmetric grouping lead to lower multipliers, as elaborated in Table 7. Our regional asymmetry does not artificially inflate the multiplier, but rather it helps us obtain a closer multiplier to the desired level.

L.3.2 A simple example

In this section, we present a simple example that captures the key reason why grouping matters for multiplier calculations. Suppose $S = \{1, 2, 3, 4\}$. We consider a strict sorting case: $A_j = j$ and $F_j = j$.

Baseline (asymmetric grouping) Define the Good group as $\{3, 4\}$ and the Poor group as $\{1, 2\}$.

$$A_G = 3.5 \tag{212}$$

$$F_G = 3.5 \tag{213}$$

$$A_P = 1.5 \tag{214}$$

$$F_P = 1.5 \tag{215}$$

This leads to the following aggregation of the group-level output:

$$(A_G F_G + A_P F_P)/2 = 7.25 \tag{216}$$

Symmetric group Define the Good group as $\{1, 4\}$ and the Poor group as

$\{2, 3\}$.

$$A_G = 2.5 \tag{217}$$

$$F_G = 2.5 \tag{218}$$

$$A_P = 2.5 \tag{219}$$

$$F_P = 2.5 \tag{220}$$

This leads to the following aggregation of the group-level output:

$$(A_G F_G + A_P F_P)/2 = 6.25 \tag{221}$$

No grouping No grouping leads to the following aggregation of the state-level output:

$$\sum A_j F_j/4 = 7.5. \tag{222}$$

The baseline provides a closer multiplier estimate, and this is due to the strict positive sorting:

$$6.25 \text{ (Symmetric group)} < 7.25 \text{ (Baseline)} < 7.5 \text{ (No grouping)} \tag{223}$$

References

ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83, 2411–2451.

AN, Z., A. KANGUR, AND C. PAPAGEORGIOU (2019): "On the Substitution of Private and Public Capital in Production," *European Economic Review*, 118, 296–311.

BAXTER, M., AND R. G. KING (1993): "Fiscal Policy in General Equilibrium," *The American Economic Review*, 83(3), 315–334.

BENNETT, J., R. KORNFELD, D. SICHEL, AND D. WASSHAUSEN (2020): "Measuring Infrastructure in the Bureau of Economic Analysis National Economic Accounts," BEA Working Paper.

BLUNDELL, R., L. PISTAFERRI, AND I. SAPORTA-EKSTEN (2016): "Consumption Inequality and Family Labor Supply," *American Economic Review*, 106, 387–435.

BUREAU OF ECONOMIC ANALYSIS (1925-2020): "NIPA Fixed Asset Table 4.1, Annual. https://apps.bea.gov/iTable/index_nipa.cfm, used June 15th, 2022," .

——— (1967-2020): "NIPA Domestic Product and Income Table 5.2.6, Annual. https://apps.bea.gov/iTable/index_nipa.cfm, used June 15th, 2022," .

CHODOROW-REICH, G. (2019): "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?," *American Economic Journal: Economic Policy*, 11, 1–34.

CHRISTIANO, L. J., M. TRABANDT, AND K. WALENTIN (2010): "DSGE Models for Monetary Policy Analysis," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3, pp. 285–367. Elsevier.

COOPER, R. W., AND J. C. HALTIWANGER (2006): "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 73, 611–633.

FALK, N., AND C. SHELTON (2018): "Fleeing a Lame Duck: Policy Uncertainty and Manufacturing Investment in US States," *American Economic Journal: Economic Policy*, 10, 135–152.

FERNÁNDEZ-VILLAVERDE, J., J. RUBIO-RAMÍREZ, AND F. SCHORFHEIDE (2016): "Solution and Estimation Methods for DSGE Models," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and H. Uhlig, vol. 2, pp. 527–724. Elsevier.

FISHBACK, P., AND V. KACHANOVSKAYA (2015): "The Multiplier for Federal Spending in the States During the Great Depression," *The Journal of Economic History*, 75(1), 125–162.

GRAVELLE, J. G. (2014): "International Corporate Tax Rate Comparisons and Policy Implications," Discussion paper, Congressional Research Service.

HAGEDORN, M., I. MANOVSKII, AND K. MITMAN (2019): "The Fiscal Multiplier," NBER Working Paper.

KHAN, A., AND J. K. THOMAS (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76, 395–436.

KIM, J.-Y. (2002): "Limited Information Likelihood and Bayesian Analysis," *Journal of Econometrics*, 107, 175–193.

KRUEGER, D., AND C. WU (2021): "Consumption Insurance against Wage Risk: Family Labor Supply and Optimal Progressive Income Taxation," *American Economic Journal: Macroeconomics*, 13, 79–113.

LEDUC, S., AND D. WILSON (2017): "Are State Governments Roadblocks to Federal Stimulus? Evidence on the Flypaper Effect of Highway Grants in the 2009 Recovery Act," *American Economic Journal: Economic Policy*, 9(2), 253–92.

LEE, H. (2025): "Striking While the Iron is Cold: Fragility after a Surge of Lumpy Investments," Working Paper.

NAKAMURA, E., AND J. STEINSSON (2018): "Identification in Macroeconomics," *Journal of Economic Perspectives*, 32, 59–86.

PERI, A., O. RACHEDI, AND I. VAROTTO (2023): "Public Investment in a Production Network: Aggregate and Sectoral Implications," Working Paper.

RAMEY, V. A. (2011): "Can Government Purchases Stimulate the Economy?," *Journal of Economic Literature*, 49(3), 673–85.

——— (2020): "The Macroeconomic Consequences of Infrastructure Investment," NBER Working Paper.