THE PUBLIC AGGLOMERATION EFFECT

Urban-Rural Divisions in Government Efficiency and Political Preferences

Theo Serlin*

October 2023

Abstract

Why and when do urban-rural divisions structure politics? The emergence of this divide in the US in the 1930s is inconsistent with canonical theories of cleavages. This paper introduces an explanation: agglomeration effects. The provision of government services should be more efficient in urban environments because of nonrivalries, economies of scale, and access costs. If the public sector in cities is more efficient, and voters face a tradeoff between taxation and government spending, urban voters should support more spending. When redistribution is salient, one should observe an urban-rural electoral divide. As predicted by a formal model, more-urban locations faced lower costs of providing public services and shifted towards the Democrats as the party implemented the New Deal. In addition, urban voters were more supportive of government spending. In the UK, the urban-rural divide also accompanied the rise of redistributive politics. Agglomeration effects influence preferences for redistribution and create political cleavages.

*Stanford University, tserlin@stanford.edu. I thank Avi Acharya, Jim Fearon, Vicky Fouka, Adriane Fresh, Judy Goldstein, Justin Grimmer, Andy Hall, Zuhad Hai, Jens Hainmueller, Rebecca Perlman, Jonathan Rodden, Ken Scheve, Denis Tchaouchev, Shun Yamaya, and participants in the Stanford Economic History and Historical Social Science Workshops, the Summer Working Group in IPE/IR, and the 2023 APSA meeting for insightful comments.

1 INTRODUCTION

Elections in the United States, and in other democracies, often pit economically conservative rural voters against economically liberal urban voters.[1] This spatial pattern is puzzling, given that rural voters are often poorer than urban voters, and might expect to benefit materially from left-wing economic policies. In countries like the US and UK where the electoral system magnifies the influence of rural voters, this division leads to unrepresentative policy outcomes (Rodden, 2019). It has not, however, always existed. Figure 1 shows the share of the two-party vote won by the Democratic Party in counties outside the former Confederacy, aggregated by counties with above- and below-average urban population shares in 1930. The urban-rural divide began, abruptly, in 1936.[2] Before that point, there is little evidence of systematic rural conservatism: in the 1880s and 1890s, the economically radical Populist Party attracted much of its support from rural areas.

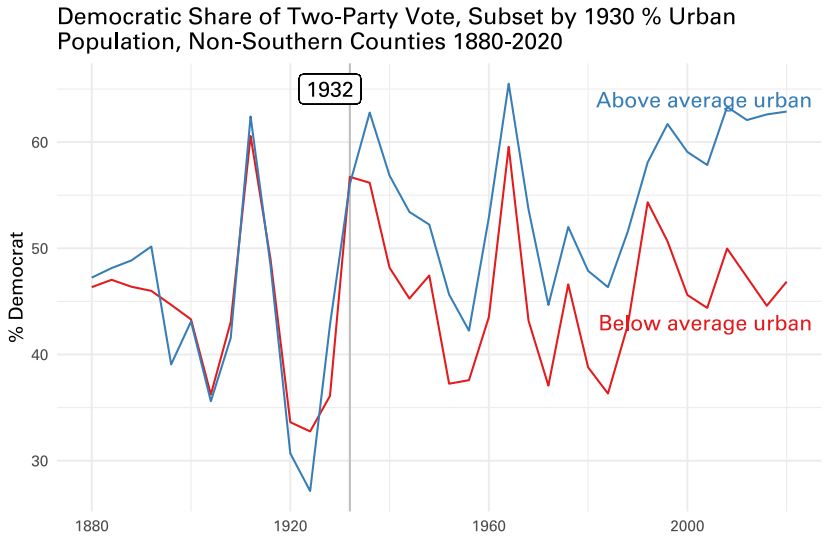

Figure 1: Emergence of the urban-rural divide after 1932

[1]Rodden (2019) documents this divide in the US, Canada, Britain, and Australia, Gimpel et al. (2020) show this discrepancy in the US cannot be attributed solely to compositional differences. In all 17 democracies for which Gethin, Martínez-Toledano and Piketty (2022) collect data on urban status and voting, rural voters are more likely to vote for right-wing parties.

[2]This divergence was not due to inter-regional realignments: regressing the county-level Democratic vote on the 1930 urban share interacted with year indicators, with state-by-year fixed effects uncovers a similar pattern (see Figure A-5).

The emergence of the urban-rural divide in 1936 is inconsistent with prominent theories of political cleavages. Lipset and Rokkan (1967) viewed the urban-rural cleavage as one of a series of longstanding social divisions created by nation-building and the industrial revolution. In Europe, these divisions provided the basis for political parties which structured politics through the 20th century.[3] This foundational account would not predict the emergence of an urban-rural cleavage in the 1930s, long after the industrial revolution. Nor would Rogowski's (1989) derivation of cleavages from the distributional effects of trade. In that account, relatively scarce factors disadvantaged by trade mobilize against relatively abundant factors. In the interwar United States, the prediction is for protectionist labor to oppose land and capital, replacing a 19th-century rural-urban cleavage between manufacturing and agriculture. While the 1930s were characterized by class politics, Figure 1 shows not the eclipse but the emergence of a rural-urban divide. Urban voters did not mobilize for higher tariffs; they elected Democratic administrations which liberalized trade policy (Hiscox, 1999). The 1936 divergence also predates the postwar racial realignment and politicization of abortion which play a prominent role in analyses of more recent urban-rural divisions (Rodden, 2019; Gimpel et al., 2020).

This article draws on economic geography to explain why and when the urban-rural divide structures politics. An extensive literature in urban and spatial economics documents the existence of agglomeration economies, whereby productivity in a given area increases with population size and density (see for instance Ciccone and Hall 1996; Bleakley and Lin 2012 and Ahlfeldt et al. 2015). This literature argues that private-sector production is more efficient in cities because of firm-level economies of scale, better matches between firms and heterogeneous workers, inputs, and consumers, and knowledge spillovers (Duranton and Puga, 2004).

These phenomena should apply all the more so to public sector provision. Public goods

[3]Scholarship building on Lipset and Rokkan (1967) emphasizes the durability of these cleavages (Bartolini and Mair, 1990), and analyzes how new cleavages emerge from broad social changes (Inglehart, 1997; Kriesi, 1998; Hooghe and Marks, 2018; Ford and Jennings, 2020).

that are not completely rivalrous are subject to economies of scale. Geographical remoteness makes it difficult for rural residents to access public services. Fixed costs of administration make the provision of programs more costly—on a per-person rate—in lower-population areas. Research in public administration finds that the provision of basic public services is more expensive in sprawling suburbs than in compact cities (Carruthers and Ulfarsson, 2003). Research in development economics finds that rural residents receive less from redistributive government programs, even when accounting for the amount disbursed to them (Olken, 2006). The point is not just that the size of jurisdictions affects efficiency, as argued by the fiscal federalism literature (Oates, 1972; Alesina and Spolaore, 2003). The need for government services to reach citizens means that transportation costs constrain economies of scale in low-density areas, regardless of the structure of government.

If government provision is less efficient in rural areas, the tradeoffs between taxation and spending faced by rural and urban voters are different. If the value of the marginal dollar of government spending is lower in rural areas than urban areas, rural voters should, all else equal, prefer less government spending and lower taxes. If the size of government is a major issue for electoral competition, voters in rural areas should be more likely than urban voters with similar incomes to support economically-conservative politicians.

That the urban-rural divide emerged in US politics when redistribution became important for national elections lends suggestive support for this thesis. Between 1932, when there was no urban-rural division in voting, and 1936, when the cleavage emerged, the Roosevelt administration expanded the size and reach of the Federal Government. Wallis (1984) documents that the federal share of government expenditure radically increased after 1932 (see Figure A-7). Public welfare and education, which had been funded almost entirely at the local level, began to be funded by the Federal Government, in part through grants to state and local governments (Wallis, 1984, 2000).[4]

The size of government became relevant to national elections between 1932 and 1936. In

[4]Figure A-8 shows the growth of intergovernmental grants in these areas after 1932.

the 1932 election, the Hoover campaign's internal polling identified the key issues dividing voters as prohibition and the tariff (Norpoth, 2019). While voters rejected the incumbent Hoover administration, the election was not a referendum on the Roosevelt administration's subsequent policies. On the campaign trail, Roosevelt prevaricated, at some points speaking of the Federal Government's responsibility for the unemployed, at others calling for fiscal retrenchment (Schlesinger, 1957, 435). Elite observers found few fundamental distinctions between the parties on economic policy (Schlesinger, 1957, 434). In contrast, Roosevelt's 1936 election victory was, in Key's (1966, 33) judgement, "a resounding ratification of the new thrust of governmental policy" in the face of Republican promises to "return to the pre-1932 status quo." Divergence over redistribution between 1932 and 1936 is also evident in the party platforms. Figure 2 plots the number of references to expanding the welfare state minus the number of references to limiting the welfare state, for the two parties, in the Manifesto Project Database (Lehman et al., 2023).

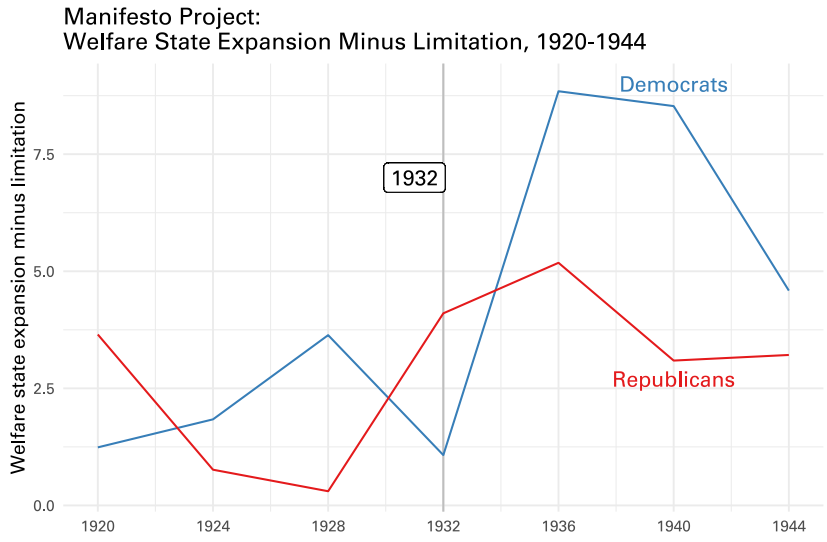

Figure 2: Divergence of parties on welfare state expansion, 1932–1936, measured in the Manifesto Project (Lehman et al., 2023)

As the parties polarized on redistribution, preferences for redistribution began to influence vote choice. Survey data shows economically-liberal voters shifting into the Democratic

coalition in 1936 and subsequent elections (Key, 1966; Caughey, Dougal and Schickler, 2020). Only after 1932 did areas that had supported left-wing third parties begin to support the Democrats (Hirano and Snyder, 2007).[5] The emergence of the urban-rural divide was contemporaneous with the divergence of the parties over the size of government.

This article develops and tests these intuitions about agglomeration, government efficiency, preferences over government spending, and urban-rural political divisions. I analyze a simple formal model that embeds agglomeration effects into a Romer-Meltzer-Richard model of redistribution. Unlike in Meltzer and Richard (1981), the government provides valuable services, and provides these services more efficiently in areas with larger populations. This agglomeration effect means that the marginal dollar of spending benefits urban voters more than their rural compatriots. The model thus predicts that urban voters should, all else equal, support higher levels of taxation and spending. Introducing a probabilistic voting framework produces the prediction that if the parties diverge on redistributive policy, urban voters should shift towards the more redistributive party, and rural voters towards the less redistributive one.

I test that prediction in three stages. First, I show that urbanization is associated with the shift towards the Democrats between 1932 and 1936. I then develop a measure of government efficiency and examine its relationship with urbanization and the New Deal realignment. Lastly, I use survey data to show that urban voters supported increased government spending and the New Deal. The formal model suggests a difference-in-differences specification, interacting time-invariant urbanization and efficiency with the post-1932 divergence between the parties on economic policy. I first use county-level data on voting, census microdata and sub-county data on population concentration from the Census Place Project (Berkes, Karger and Nencka, 2022) to analyze the relationship between different measures of urbanization and the 1932–1936 shift in voting. I focus on the urban share of the population and a measure of population

[5]Figure A-6 confirms this pattern: county level Socialist Party vote share over the period 1904–1920 was uncorrelated with the Democrats' share of the two-party vote in 1932, from 1936 onwards the two were strongly and positively correlated.

agglomeration suggested by the model. Going from the 25th to the 75th percentile of the urban share is associated with a 5 percentage point increase in the Democratic vote between 1932 and 1936 in the least restrictive specification, and a 1.5 percentage point change in the most. These results are robust to controlling for the presence of manufacturing, agriculture, ethnic minorities, and union-intensive industries.

Second, I use data on pre-New Deal government spending to develop a measure of efficiency based on the share of spending allocated to voter-facing services like schools, libraries, and police, relative to government overhead, and show that this variable is associated with the 1932–1936 realignment. Local government efficiency ought to affect vote choice in national elections because many New Deal programs were implemented in part through local governments (Wallis, 1984, 2000; Fetter, 2017). I digitize records of government spending across categories in 1932 for non-Southern towns with more than 8,000 residents. This measure is positively correlated with the urban population share.[6] Controlling for the urban population share, this efficiency measure is orthogonal to economic and social variables bundled with urbanization, like the presence of agriculture or manufacturing, but correlated with the theoretically-consistent measure of agglomeration. Counties with more efficient governments shifted towards the Democrats between 1932 and 1936. This result holds when controlling for the urban population share, which provides evidence of a shift due to efficiency not attributable to other factors correlated with urbanization.

Third, I use individual-level survey data from the period to show that urban voters were more supportive of higher taxes, more government spending, and New Deal programs. This evidence supports the direct predictions of the model about an urban-rural divide in policy preferences, and makes it more plausible that the voting results are due to the New Deal and preferences over the size of government.

While the primary case study in this paper is the New Deal, the theory should not just apply to the Democratic-Republican divide in US elections. I examine changing voting

[6] I also document a number of other margins on which government provision was more efficient in urban areas.

alignments in British elections. At the turn of the twentieth century, the area and sparsity of a constituency was uncorrelated with its support for the Conservative Party. The emergence of redistribution as an issue in British politics, first due to the Liberal Party's championing of a nascent welfare state in the 1900s, and then more strongly with the rise of the explicitly socialist Labour Party, precipitated an urban-rural divide. This relationship also applies to the emergence of the urban-rural divide in Canada in the 1960s.

This study makes three contributions. First it offers an explanation of the rural-urban electoral divide that is distinct from, though not incompatible with, existing explanations. Since Lipset and Rokkan (1967), an enormous literature across fields of political science and related social sciences has sought to explain why particular social divisions structure politics (see for instance Rogowski 1989; Bartolini and Mair 1990; Kriesi 1998; Manza and Brooks 1999; Hiscox 2002; Posner 2004; Caramani 2004; Hooghe and Marks 2018; and Gethin, Martínez-Toledano and Piketty 2022). In Rodden's (2019) account, the emergence of the urban-rural divide in the US was a byproduct of concentrations of union labor in cities, itself a byproduct of the clustering of manufacturing in urban cores in this era.[7] My empirical results are robust to controlling for manufacturing and union-intensive industries, suggesting that there were reasons for cities to shift towards the Democrats distinct from industrial unions. Furthermore, agglomeration effects explain why often-poorer voters in rural areas did not join the Democratic coalition with the same fervor as urban workers. The success of earlier rural radicals, such as the populists, suggests that rural voters were capable of mobilizing in support of government intervention in the economy. Those earlier causes were not focused around redistribution, but around monetary policy and the regulation of banks and railroads, issues not affected by local government provision.[8]

[7]Ogorzalek (2018) also notes that cities may have been more supportive of the New Deal because of unemployment related to manufacturing.

[8]A related literature on rural support for the far-right in interwar Europe, especially in Germany, highlights a different set of mechanisms which are unlikely to apply to the US case. Moore (1966) argues the ability of powerful landowners to manipulate peasants accounts for this phenomenon. In Luebbert's (1991) account, wealthier peasants sided with the far-right because socialists threatened rural hierarchies. The US lacked powerful landowning elites outside the South, and, thanks to the influence of Southern Democrats, the New-Deal era Democratic Party if anything reinforced rural hierarchies, for instance with farm policies like

Dasgupta and Ramirez (2020) offer an explanation of rural conservatism on the Great Plains in which postwar technological developments magnified the local influence of the wealthiest industrialized farmers. This study examines a slightly earlier period, before the rise of agribusiness, and uses a different theory to explain different variation in voting.

This study is closely related to Cramer's (2016) analysis of rural consciousness. Cramer argues that rural conservatism is rooted in the belief that rural areas do not receive their fair share of resources. Cramer's interviewees argue "the government must be mishandling my hard-earned dollars, because my taxes keep going up and clearly they are not coming back to benefit people like me. So why would I want an expansion of government?" (240). This paper examines the theoretical and empirical foundations of that perception, and argues that, in addition to the social identity mechanism Cramer explores, a simpler mechanism based around government efficiency contributes to the rural-urban divide.

Second, this study introduces a new variable that explains attitudes to the size of government. These preferences are important for theories of regime type (Boix, 2003; Acemoglu and Robinson, 2006) and international openness (Ruggie, 1982) in addition to redistribution. The extensive literature on this topic has considered the importance of income (Romer, 1975; Meltzer and Richard, 1981), social mobility (Piketty, 1995; Benabou and Ok, 2001), labor market risks (Moene and Wallerstein, 2001; Alt and Iversen, 2017), racial and ethnic divisions (Alesina, Baqir and Easterly, 1999; Lee and Roemer, 2006), religion (Scheve and Stasavage, 2006; Huber and Stanig, 2011), and social identity (Shayo, 2009). Much of this scholarship considers how the inefficiency of redistribution—in the case of Meltzer and Richard (1981), due to labor market distortions—affects preferences for government spending. This study focuses on variation in the efficiency of public goods provision across space, and relates to Rueda and Stegmueller (2016) and Sands (2017)—which argue respectively that local experiences of crime and inequality influence support for redistribution—in emphasizing the importance of place-based factors for individual preferences over the size of government.

the Agricultural Adjustment Act that benefitted landowners over tenants.

Third, this article introduces a mechanism through which a person's geographical location influences their policy preferences. In focusing on how place conditions the effectiveness of different policies, this article differs from the vast literatures on sociotropism and intergroup contact. In those literatures, place influences preferences through information—voters observe the local effects of policies, whether on unemployment or the presence of out-groups—and other-regarding preferences (Kinder and Kiewiet, 1981; Mansfield and Mutz, 2009; Sands, 2017; Enos, 2016; Paluck, Green and Green, 2019). The present analysis is closer to scholarship on the politics of trade, which focuses on the local labor market effects of economic policies (Rickard, 2020; Broz, Frieden and Weymouth, 2021).

Two related studies are Ogorzalek (2018) and Lizzeri and Persico (2004). The former analyzes coordination by mayors and members of congress to lobby for urban-focused policies in the New Deal era. This article focuses on the shift of urban voters into the Democratic coalition, and complements that book's focus on elites. It also presents evidence of greater efficiency of government provision in cities that provides a rationale for urban support of redistributive government spending beyond specifically urban-focused policies like slum clearance. Lizzeri and Persico (2004) offer an explanation for franchise extension. Expanding the franchise shifts electioneering from clientelism to programmatic competition. Lizzeri and Persico argue that in 19th century Britain, urbanization created new demands for public goods, especially public health infrastructure. This article focuses on a different mechanism, in which urbanization increases the efficiency of government provision. This efficiency mechanism is relevant to a broader range of government spending categories and thus can explain attitudes to redistributive government programs administered at the national level.

The rest of the paper proceeds as follows. The next section sets out the model and derives predictions about individual preferences over government spending and electoral realignments. Section 3 tests the main prediction about realignment. Section 4 presents individual-level survey evidence on the link between urbanization and support for redistribution in this period. Section 5 examines the British case. Section 6 concludes with implications for scholarship on

the politics of inequality and trade.

2 A MODEL OF REDISTRIBUTION AND POLITICAL CLEAVAGES WITH AGGLOMERATION EFFECTS

The government taxes voters at a linear rate t and allocates public spending to localities proportional to population. The pre-tax income of individual i is given by y_i, and average income is given by $\bar{y}$. In each locality l, the government provides public goods and services subject to the following locality-specific production function:

$$G_l = A_l g^{\varphi} \tag{1}$$

where g is per-capita spending, $\varphi \in (0, 1)$ is a technology parameter that captures the concavity of the production function, and A_l is location-specific productivity. Public sector productivity is subject to an agglomeration effect:

$$A_l = \bar{A} n_l^{\alpha},\ \alpha > 0$$

where n_l is the population of location l. I thus assume that in more populous locations, the government can provide more services per dollar of spending. This reduced-form approach to modeling agglomeration effects follows the economic geography literature (Allen and Donaldson, 2020). Below I set out a number of different theoretical and empirical microfoundations—economies of scale, nonrivalries, and spatial frictions to access—for this assumption and functional form.

ECONOMIES OF SCALE One way in which place size should affect the efficiency of government is through economies of scale in administrative costs. More formally, suppose that in order to spend g dollars per capita on service provision, locality l must spend $c_l g$ dollars, where $c_l > 1$ includes per-unit administrative costs. The per-capita administrative cost c_l is

decreasing in population: $c_l = \bar{c} n_l^{-\beta}$. The parameter β can be thought of as capturing the relative importance of fixed and variable costs: if $\beta = 1$, administrative costs are entirely fixed, if $\beta = 0$, they are entirely variable. If a locality receives g income per capita to spend, it can provide $g' = \frac{g n_l^{\beta}}{\bar{c}}$ in service spending per capita. The actual provision of services is subject to the production function $G_l = Z\,(g')^{\varphi}$, as in (1), where Z is productivity. Substituting in the identity for g' gives the following identity:

$$G_l = Z \left(\frac{g n_l^{\beta}}{\bar{c}} \right)^{\varphi} = Z_l \left(\bar{c}\right)^{-\varphi} n_l^{\beta\varphi} g^{\varphi}$$

If $\alpha = \beta\varphi$ and $\bar{A} = Z\bar{c}^{-\varphi}$, this setup is isomorphic to $G_l = \bar{A} n_l^{\alpha} g^{\varphi}$.

NONRIVALRIES The inherent nonrivalry of public goods may also create economies of scale. Suppose that a locality receives $n_l g$ dollars to spend on public provision, which it can spend on providing a rivalrous good and a non-rivalrous good. Citizens' preferences over the two public goods are Cobb-Douglas, $G_l = q_{l1}^{\beta} q_{l2}^{1-\beta}$, where q_{l1} is the number of units of the nonrivalrous good provided to a citizen, for which total production is all that matters, q_{l2} is the number of units of the rivalrous good, of which each citizen in l receives an equal share of the total produced, and β is the weight attached to nonrivalrous goods. If the municipality produces each service with the production function $q_{lj} = Z g_{lj}^{\varphi}$, and spends $\bar{g}_l$ on the nonrivalrous good and the remainder on the other good, then citizen utility from government provision is

$$G_l = Z\,(\bar{g}_l)^{\varphi} \left(\frac{(n_l g - \bar{g}_l)^{\varphi}}{n_l} \right)^{1-\beta}$$

Maximizing this function implies that the government spends fraction β on the nonrivalrous good, giving utility of:

$$G_l = Z n_l^{\beta+\varphi-1} \left(g \beta^{\beta} (1-\beta)^{1-\beta} \right)^{\varphi}$$

Thus for $\alpha = \beta + \varphi - 1$, and $\bar{A} = Z \left(\beta^{\beta} (1-\beta)^{1-\beta} \right)^{\varphi}$ this setup is isomorphic to $G_l = \bar{A} n_l^{\alpha} g^{\varphi}$.

ACCESS COSTS Geographic barriers to accessing government services may also make provision more efficient in more populous areas. Suppose that a government program can only be accessed in-person, and government in a locality has amount $n_l g$ to spend to provide q_l service locations, which are produced with the following technology:

$$q_l = Z\left(n_l g\right)^{\varphi}$$

The number of service locations is an increasing function of total spending ($n_l g$). In the locality, population is uniformly distributed on the $[0, 1]$ interval, and so population density at any point is n_l. Only citizens within $\frac{k}{2}$ distance from a service location access the good, but access is subject to congestion costs, such that if m people are close enough to access a service station, only m^{ν} will receive the service, where $\nu < 1$ determines how sharply these congestion costs increase with density. Accessing the service gives utility γ. The ex-ante probability of a citizen accessing the good is thus $k^{\nu} n_l^{\nu-1} q_l$, and the ex-ante expected utility of a citizen from government spending is

$$G_l = \gamma k^{\nu} Z\left(n_l\right)^{\varphi+\nu-1} g^{\varphi}$$

If $\alpha = \varphi + \nu - 1$ and $\bar{A} = \gamma k^{\nu} Z$, this setup is also isomorphic to $G_l = \bar{A} n_l^{\alpha} g^{\varphi}$.

2.1 *Preferences over the Size of Government*

Individual utility is linear and additive in government services and post-tax income: $u_{il} = G_l + (1-t) y_i$. As spending is proportional to population, per-capita spending in each location is average income $\bar{y}$ multiplied by the tax rate t. Substituting in these various identities gives the following expression for utility as a function of pre-tax income, taxation, and agglomeration:

$$u_{il} = \underbrace{\bar{A} n_l^{\alpha} \left(t\bar{y}\right)^{\varphi}}_{\text{public good}} + \underbrace{(1-t) y_i}_{\text{post-tax income}}$$

The term on the left is increasing in t—a higher tax rate means more government revenue and more public services—and the term on the right is decreasing in t, as a higher tax rate means less post-tax income. A voter's preferred tax rate depends on this tradeoff between private and public benefits. As this expression is concave in t, taking first order conditions gives individual i's optimal tax rate (if less than one) and level of government spending:

$$t^*_{il} = \left(\frac{\bar{A} n_l^{\alpha} \bar{y}^{\varphi}}{y_i}\right)^{\frac{1}{1-\varphi}}, \; g^*_{il} = \left(\frac{\bar{A} n_l^{\alpha} \bar{y}}{y_i}\right)^{\frac{1}{1-\varphi}} \tag{2}$$

Individuals favor more government when they are poorer relative to the average ($\frac{\bar{y}}{y_i}$ is large), and when they are in more populous locations (n_l is large).

2.2 *Urban-Rural Electoral Divisions*

These results suggest an urban-rural division in preferences over taxation policy. I embed these preferences in a probabilistic voting framework (Lindbeck and Weibull, 1987), to derive predictions about rural-urban voting patterns depending on whether the parties are divided over spending policy.

There are two parties, L and R, which propose tax rates, $t^L \geq t^R$. Voters choose which to vote for based on both the proposed tax rates and place- and individual-level preference shocks. Individual i in location l votes for party L if the following inequality is satisfied:

$$\underbrace{\bar{A} n_l^{\alpha} \left(t^L \bar{y}\right)^{\varphi} + \left(1 - t^L\right) y_i}_{\text{welfare under } t^L} + \underbrace{\psi_l - \varepsilon_i}_{\text{bias}} > \underbrace{\bar{A} n_l^{\alpha} \left(t^R \bar{y}\right)^{\varphi} + \left(1 - t^R\right) y_i}_{\text{welfare under } t^R}$$

The term ψ_l captures place-specific factors that influence her preference for one party, and ε_i captures idiosyncratic deviations from the location mean, which are drawn identically and independently across individuals from a uniform distribution over the interval $\left[-\frac{1}{2\theta}, \frac{1}{2\theta}\right]$. The parameter θ governs how responsive vote choices are to differences in economic policy. The

probability of i in l voting for L is

$$\begin{aligned} s_{il} &= P\left(\varepsilon_i < \bar{A} n_l^{\alpha}\left(\left(t^L \bar{y}\right)^{\varphi} - \left(t^R \bar{y}\right)^{\varphi}\right) - \left(t^L - t^R\right) y_i + \psi_l\right) \\ &= \theta\left(\bar{A} n_l^{\alpha}\left(\left(t^L \bar{y}\right)^{\varphi} - \left(t^R \bar{y}\right)^{\varphi}\right) - \left(t^L - t^R\right) y_i + \psi_l\right) + \frac{1}{2} \end{aligned} \tag{3}$$

This expression implies that, if $t^L > t^R$, holding all else constant, residents of more populous locations are more likely to vote for the high-tax party, and that higher-income individuals are less likely to vote for the high-tax party.[9] This pair of results echoes Gelman et al.'s (2010) finding that while poorer, more rural states tend to vote Republican and richer states Democratic, within both types richer voters are more likely than poorer voters to vote Republican.[10] Furthermore, increasing $t^L - t^R$ will increase the divergence in voting between urban and rural locations, as well as between high- and low-income individuals.

The precise mechanism that causes the parties to diverge on policy is unimportant for the paper's focus on how policy divergence generates an urban-rural cleavage. A model of policy choice in this context would need to generate both platform convergence and divergence as equilibrium outcomes. Appendix A sets out one model that generates that prediction. The key assumption is that a party chooses the policy preferred by the median voter within that party. If the parties converge on policy, they both win similar numbers of votes from rural and urban voters, but if they diverge, the party choosing the higher tax rate wins more votes from urban voters.

The model predicts that when the parties are divided on issues of redistribution, urban voters should, all else equal, support the pro-redistribution party. The reason for this divergence is that government provision is more efficient in cities. Because government provision is more efficient there, urban voters support higher rates of taxation and government spending. The

[9]While I simply assume that parties allocate spending to locations equally, all else equal, voters in more populous locations will be more likely to support the party that promises to spend more in that location.

[10]That wealthier residents of Blue states tend to live in more urban environments, like Manhattan, than their peers in Red states rationalizes Gelman et al.'s (2010) other result of a steeper gradient between income and voting Republican in Red states.

remainder of the paper tests these predictions.

3 EXAMINING THE EFFECTS OF AGGLOMERATION ON THE 1932–1936 REALIGNMENT

I use the divergence of the parties on redistributive policy between 1932 and 1936 to test the prediction of Equation (3). Aggregating Equation (3) at the county level gives the following expression for the county-level vote for the higher-tax party:

$$s_c = \underbrace{\theta\left(\left(t^L\bar{y}\right)^{\varphi} - \left(t^R\bar{y}\right)^{\varphi}\right)}_{\text{Weakly-positive scalar}} \underbrace{A_c}_{\text{Efficiency}} - \underbrace{\theta\left(t^L - t^R\right)y_c}_{\text{Effect of income}} + \underbrace{\theta\psi_c + \frac{1}{2}}_{\text{County-specific bias}} \tag{4}$$

where s_c is vote-share for the left-wing party in county c, y_c is average income in county c, ψ_c is the population-weighted average of place-specific biases ψ_l, and A_c is population-weighted average productivity:

$$A_c = \frac{\sum_{l\in c} n_l\left(\bar{A}n_l^{\alpha}\right)}{\sum_{l\in c} n_l} = \bar{A}\left(n^{\alpha}\right)_c \tag{5}$$

The term in front of A_c in (4) is a weakly-positive scalar for $t^L \geq t^R$, strictly positive if $t^L > t^R$. (4) implies a linear relationship between efficiency and voting Democrat. Estimating this relationship by OLS is complicated by the concern that county-specific biases ψ_c might be correlated with efficiency A_c.

The sharp divergence on policy between 1932 and 1936 provides a way to address that concern. If the parties converge on the issue of spending, that is, if $t^R = t^L$, then (4) reduces to $\theta\psi_c + \frac{1}{2}$; electoral outcomes are determined by issues other than spending. As discussed above, in 1932 that appeared to be the case, but by 1936 the parties had diverged. Adding time subscripts to (4) and differencing between these two periods gives the following expression for the change in Democratic vote share:

$$s_{c,36} - s_{c,32} = \underbrace{\theta\left(\left(t^L_{36}\bar{y}_{36}\right)^{\varphi} - \left(t^R_{36}\bar{y}_{36}\right)^{\varphi}\right)}_{\text{Strictly-positive scalar}} \underbrace{A_{c,36}}_{\text{Efficiency}} - \underbrace{\theta\left(t^L_{36} - t^R_{36}\right)y_{c,36}}_{\text{Effect of income}} + \underbrace{\theta\left(\psi_{c,36} - \psi_{c,32}\right)}_{\text{Change in bias}} \tag{6}$$

Holding income and changes in non-economic preferences for the parties constant, counties with more productive governments should have shifted towards the Democrats between 1932 and 1936. Differencing relaxes the identification concern by requiring only that changes in non-economic preferences for the Democrats between 1932 and 1936 are uncorrelated with levels of efficiency.

3.1 *Estimating Equation and Data*

Equation (6) suggests the following estimating equation:

$$\Delta \%Democrat_c = \beta Efficiency_c + \mathbf{x}_c'\gamma + \delta_{s(c)} + \varepsilon_c$$

where $\Delta \%Democrat_c$ is the change in the Democrats' share of the two-party vote in county c between 1932 and 1936, $\mathbf{x}_c$ is a vector of controls, $\delta_{s(c)}$ is a state fixed effect, which given the differenced specification is equivalent to a state-by-year fixed effect, ε_c is an error term, and $Efficiency_c$ is a measure of agglomeration-related efficiency, discussed below. I restrict the analysis, except where specified, to counties outside the former Confederacy, because I would not expect to observe much variation in voting in the South given the extremely restrictive franchise and lack of partisan competition in this period. I estimate these models by OLS with standard errors clustered at the state level to account for spatial autocorrelation. I use data on voting from Clubb, Flanigan and Zingale (2006).

I focus initially on two measures of agglomeration-related efficiency. First, I use the urban share of the population, measured using the 1930 census microdata (Ruggles et al., 2021). The census classified those living in places with more than 2,500 residents as urban. This metric should not capture all the effects of place size and population density on public sector efficiency and thus on voter preferences. It is however important to establish that the divergence shown in Figure 1 was an urban-rural realignment.

Second, I calculate the empirical analogue of Equation (5), the population-weighted

average of sub-county place populations raised to the agglomeration elasticity α, which I set at 0.078 using an estimate based on the cost of police and fire services from Appendix B. I use place-specific populations in 1930 from the Census Place Project (Berkes, Karger and Nencka, 2022), which uses the information on locations recorded in the census microdata to estimate sub-county geography. I refer to this variable as agglomeration and winsorize it at the 2.5th and 97.5th percentiles to ensure that my results are not driven by outliers.[11] Note that regressions of the change in voting between 1932 and 1936 on this measure are almost exact empirical analogues of Equation (6).

My baseline set of controls are the shares employed in agriculture and in manufacturing, the white and foreign-born shares of the population, measured in the 1930 census microdata, and union potential, which I calculate following Collins and Niemesh (2019) by interacting unionization rates at the industry level in 1939 from Troy (1957) with industry shares in the 1930 census microdata. While there is no county-level data on union membership, union potential should capture exposure to unionization based on the industry mix. To address the concern that voters in high-efficiency counties differed in income, in an additional robustness check I also control for levels and changes in retail sales per capita and bank deposits per capita using data from Fishback, Horrace and Kantor (2005) and the FDIC (Manson et al., 2020), respectively. In another robustness check related to unionization I control for the number of locals of and strikes by the Industrial Workers of the World relative to population. Other robustness checks control for farms and farm output per capita from the Census of Agriculture, and the population share unemployed in 1930 and the illiteracy rate among those over the age of 10, from the Census of Population (Manson et al., 2020).

3.2 *Results for Voting*

Table 1 shows the results of regressions of the change in voting Democrat after 1932 on these measures of urbanization. Model (1) shows the relationship between the 1930 urban share

[11] Neither the substantive nor the statistical significance of my results is sensitive to this decision (Table A-10).

	Δ% Democrat 1932–1936			
	(1)	(2)	(3)	(4)
% urban	0.112***	0.035***		
	(0.008)	(0.011)		
agglomeration			0.183***	0.064***
			(0.014)	(0.019)
Controls		x		x
DV mean	-0.011	-0.011	-0.011	-0.011
N	1761	1761	1754	1754
R^2	0.529	0.582	0.531	0.581

This table shows the results of regressions of the county-level change in the Democrats' share of the two-party vote between 1932 and 1936 against the urban share of the population in 1930, and agglomeration, calculated as the population-weighted average of place size raised to the power $\alpha = 0.078$. All models include state fixed effects. Even-numbered models also control for the shares employed in agriculture and manufacturing, the white and immigrant shares of the population, and union potential. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 1: Effects of urbanization on realignment towards the Democrats after 1932

and the change in votes, and suggests that a 1 percentage point increase in the urbanized share was associated with around a 0.1 percentage point increase in the Democrats' share of the two-party vote. Model (2) adds baseline controls for the economic and ethnic composition of counties. With these controls the coefficient on urbanization is around a quarter of the size, but still positive and significant, suggesting that there was an effect of urbanization on voting after 1932 distinct from the effects of agriculture, manufacturing, immigration, or race. The interquartile range of urbanization is around 44 percentage points, the effect on the 1932–1936 realignment of going from the 25th to the 75th percentile is thus around 5 percentage points in the base specification and 1.5 percentage points with controls. Models (3) and (4) show larger positive effects for the theoretically-consistent agglomeration metric—the standard deviations of these two independent variables are similar. This difference in magnitudes makes sense given that the agglomeration variable should more precisely capture variation in government efficiency that should affect the realignment.

3.3 *Threats to Identification*

The differenced specification should account for time-invariant non-economic factors that might influence voting. There are two key threats to identification. First, as in a standard difference-in-differences design, there is a concern that locations with higher rates of efficiency would have followed different trajectories to low-efficiency locations even in the absence of the parties' divergence. I address this concern with standard difference-in-differences checks, verifying that high-efficiency locations were not following differential trends up until 1932, in Figure A-11.

Second, other variables correlated with efficiency might influence the change in voting between 1932 and 1936. These factors could either relate to modeled levels of income (the second term in Equation 6) or to un-modeled changes in voters' non-economic preference for the Democratic Party (the third term). The most pressing concern is that other factors correlated with urbanization interacted with the New Deal policy divergence to influence voting for the Democrats. The main alternative explanations are the realignments of urban Black and immigrant voters into the Democratic Party, and the embrace of the Democrats by labor unions, the membership of which tended to be urban. These explanations are broadly consistent with the theoretical model. Urban minority voters and unionists may have joined the Democratic coalition precisely because, due to their location, they stood to benefit most from the redistributive state.[12] Yet it is still a concern that factors unrelated to agglomeration—for instance the changing positions of the parties on civil or labor rights—accounted for these specific realignments.

The timing of the rural-urban split between 1932 and 1936 is somewhat inconsistent with explanations based around race, immigrants, or unions. Schickler (2016), who dates

[12]Table A-8 shows that urban areas did not receive more New Deal funds per capita, using data from Fishback, Kantor and Wallis (2003), ruling out an explanation in which the targeting of New Deal funds to urban constituencies created the divergence. Major New Deal programs like the Agricultural Adjustment Administration were targeted at the agricultural sector, a pattern attributable to the influence of Southern Democrats in Congress and to the dire state of agriculture in this period. Scholarship on the New Deal emphasizes the importance of both economic hardship related to the Depression and political considerations for the allocation of New Deal funds (Fishback, Kantor and Wallis, 2003).

the civil rights realignment earlier than most scholars (Carmines and Stimson, 1989), argues that it was only in the late 1930s that racial equality became a goal of the New Deal and that African Americans became disproportionately supportive of the Democrats in the 1940s. Furthermore, the 1932–1936 realignment predates the second Great Migration; the northern Black population was at the time relatively small and unlikely to totally account for the emerging urban-rural divide. Analyses of the entry of various "white ethnic" immigrant groups into the Democratic Party generally focus on the critical election of 1928, when the party nominated the Irish Catholic Al Smith, instead of 1936 (Key, 1955). Thus even if immigrants continued to shift into the Democratic Party between 1932 and 1936, the absence of differential trends in voting by high-agglomeration locations would constitute evidence against an immigrant realignment driving the result. The 1936 election occurred just as the mid-20th century surge in union membership was beginning (see Figure A-10). An alternative mechanism in which unions mobilized urban workers into the Democratic Party would predict a much more gradual emergence of the urban-rural divide over the 1940s and 1950s than that shown in Figure 1.

Controlling for the shares employed in manufacturing and agriculture, the white and immigrant shares of the population, and union potential provides evidence against simple explanations for Figure 1 in which, for instance, the presence of manufacturing entirely explains the emergent urban-rural divide.[13] Examining the dynamic effects of these controls on voting, in the same event-study specification used to study the dynamic effects of the urban share, provides further evidence against that idea. Figure A-12 shows the coefficients from all these variables interacted with year indicators, controlling for the other variables interacted with year indicators and for county and state-by-year fixed effects. The control variables are all correlated with differential trends in voting prior to 1932, and do not show an obvious 1932–1936 divergence. Indeed, these differential trends fit with what one would

[13]Table A-7 shows these results are also robust to controlling for various proxies for income in 1932, the change in income between 1932 and 1936, and union density, as well as for other measures of agriculture and economic development, and for the change in voting between 1928 and 1932.

expect given the literature, for instance, the relationship between the immigrant share of the population and voting shifts noticeably in 1928. It seems these factors influenced voting through long-run shifts and not through specific alternative mechanisms interacting with the policies implemented by the first FDR administration.

3.4 *Direct Effects of Efficiency*

While these results show that more urban areas shifted towards the Democrats after 1932, they do not necessarily show that government efficiency was a driver of this shift. In addition, though the addition of controls suggests that residual variation in urbanization not due to industry or ethnic mix accounts for some of the shift, one might still be concerned that other correlates of urbanization account for that result. I address both concerns by developing a measure of efficiency from government spending across categories, and studying its effects on voting.

The *Financial Statistics of State and Local Government: 1932* provide information on county and sub-county spending on a host of different categories. Derenoncourt (2022) digitized the county records for non-Southern states. I digitize the records for non-Southern towns with more than 8,000 residents, and the county records for West Virginia.[14] Using data on pre-New Deal local government efficiency is especially appropriate given that many New Deal programs were in part implemented through local governments (Fetter, 2017). Referring back to the derivation of economies of scale in Section 2, I classify spending on policing and firefighting, public health, charity and poor relief, education, libraries, recreation, and development as spending on goods received by citizens, g. I add general government expenses and miscellaneous expenses to compute total government spending $c_l g$.[15] Taking the ratio of these two gives a measure of government efficiency $1/c_l$, the share of spending going on services that directly benefit voters. I winsorize this variable and refer to it as *overhead*

[14]The data source reports information for towns with 8,000–30,000 and more than 30,000 residents separately from towns with 2,500–8,000 residents.

[15]The one other category is highway maintenance, which I use in Appendix B.

efficiency.

This is a narrow measure of government efficiency, specifically relating to administrative costs in local government, and should not capture all forms of government efficiency. It is however correlated with higher rates of school attendance and lower infant mortality (Table A-4), variables considered by public policy scholars in this period to be outputs of good government (Walker, 1930). It is negatively correlated with government overhead spending per capita, but orthogonal to non-government spending (Figure A-4), suggesting that the measure picks up economies of scale in overhead costs, not more expansive government in general. Overhead efficiency was also negatively correlated with the cost of administering unemployment relief, suggesting the variable plausibly captures variation in government efficiency that affected redistributive programs. In Michigan, the ratio of administrative costs to total disbursements from the state's emergency relief programs, which I digitize from Haber (1935), was negatively correlated with overhead efficiency (Table A-5).

Experts in the 1930s recommended the use of general government expenses relative to other spending to gauge government efficiency. Galbraith (1934, 56), analyzing counties in California, noted that general government expenses "are not a part of the direct services which counties exist to perform. ... many of the expenditures for general government are expenditures for the *overhead* costs of the county." Because such spending only indirectly facilitated "direct services" like policing, education, and poor relief, their relative share of spending provided a measure of the extent to which administration crowded out service provision: "the volume of general government expenditure—for assessment, collection of taxes, auditing, etc.—in relation to other expenditures [can] be recognized as one of the valuable indicators of the efficiency of county operation." Galbraith also noted that there was little demand for improved service quality in the case of government administration: "a campaign for ... more or better assessing, tax collecting, or auditing, has small popular appeal indeed" (20). This intuition is supported by polling data from the era. In one 1937 poll, government running expenses were the least popular category of public spending. 67% of non-Southern

respondents stated they wanted the government to decrease running expenses; 6% supported an increase.[16]

The town-level data shows clear evidence of agglomeration economies. Figure 3 documents a positive linear relationship between the log of town population and log overhead efficiency, as predicted by the expression $\frac{1}{c_l} = \frac{n_l^\beta}{\bar{c}}$. Estimating this relationship by OLS suggests $\beta = 0.05$ (Table A-1). The positive association between population and administrative efficiency is robust to the addition of state fixed effects.

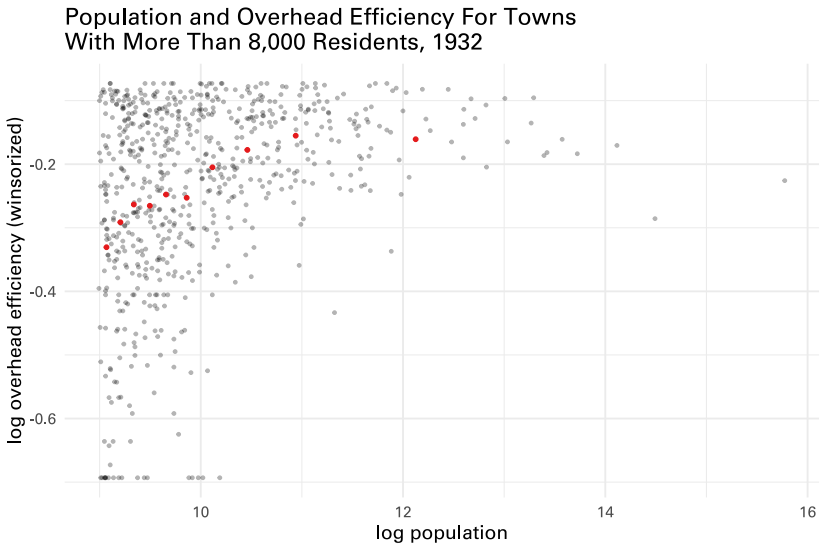

Figure 3: In larger towns, a larger share of government spending went on voter-facing services relative to administrative expenses

This figure shows the relationship between town population and overhead efficiency, for non-Southern towns with more than 8,000 residents in 1932 reporting government spending by category

Voters in the 1930s were aware of these patterns. A 1937 treatise on rural public administration observed that "a visit to the various offices of a typical rural county will often afford a demonstration of the backwardness of such areas in adopting new methods of office practice" (Lancaster, 1937, 121). Rural bureaucracy was inefficient in part because the small scale prevented specialization and the adoption of "modern scientific administration" (115). The same issue of scale prevented rural areas from providing adequate services. Anderson (1936, 29) noted that "a community of 5,000 persons ... could not conveniently support a complete

[16]Gallup Poll # 1937-0101, October 20-25 1937.

public health unit." Urban density was essential for efficient public services: "Obviously, any city whose population spreads thinly over a large area will either have to deny some of its inhabitants some of the normal municipal services, or else it will have higher per capita expense" (32). Agglomeration effects were most evident in the condition of rural schools: "The expense of teaching is high, no matter how poorly the teacher is paid, and the quality of work done is usually poor" (33).

I study the effects of overhead efficiency on the 1932–1936 change in the Democratic vote, with and without controls for urbanization. These models both directly test the effects of one measure of efficiency on the realignment, and—in the models controlling for urbanization—make it more plausible that this effect is not simply due to factors bundled with urbanization. The logic of the latter approach is that the residual from regressing overhead efficiency on the urban share should be correlated with more theoretically-motivated forms of agglomeration and should still affect changes in voting over the 1932–1936 period through the efficiency mechanism. It should not however be correlated with variables that are bundled with urbanization. Figure 4 shows point estimates and 95% confidence intervals from regressions of outcome variables against overhead efficiency, with and without controls for the urban share. This efficiency metric is strongly correlated with the urban share and with variables correlated with urbanization in this period, such as manufacturing, agriculture, and immigrants. However, controlling for the urban share, overhead efficiency is orthogonal to these variables. This exercise raises confidence that a regression of changes in voting on overhead efficiency, controlling for the urban share, should not pick up the effects of unobservables correlated with urbanization. Controlling for the urban share, overhead efficiency is still correlated with the agglomeration metric and log population density, suggesting that the residual variation is due in part to more theoretically-consistent forms of agglomeration than the urban population share. Figure A-14 shows the spatial distribution of overhead efficiency, with and without controls for urbanization.

Table 2 shows the effects of overhead efficiency on voting. Model (1) does not include

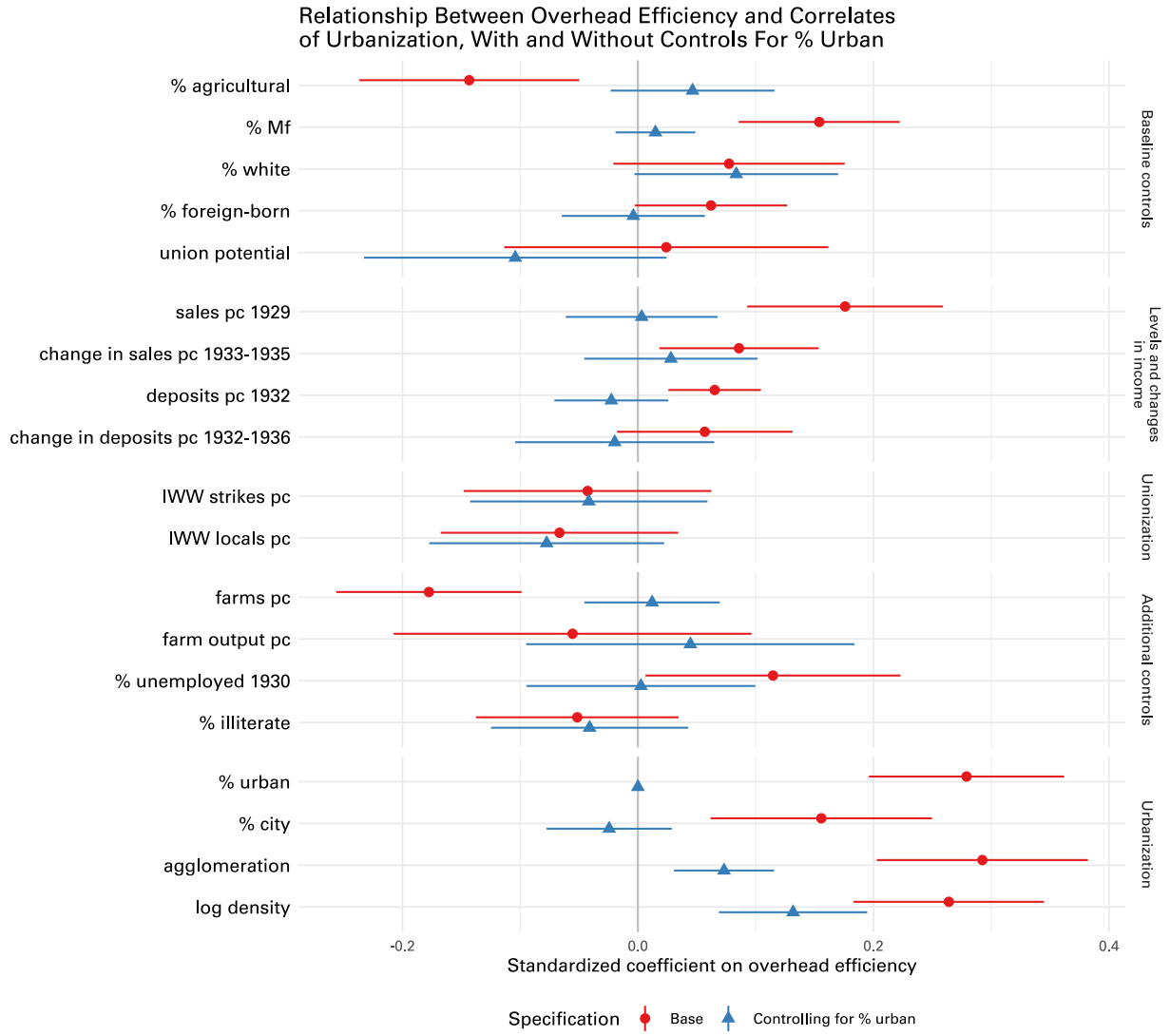

Figure 4: Controlling for % urban, the goverment provision share is uncorrelated with factors associated with urbanization and the New Deal realignment, but correlated with model-suggested measures of agglomeration

This figure shows coefficients and 95% confidence intervals from regressions on overhead efficiency. The independent and dependent variables are standardized. Red circles are from models controlling for state fixed effects, blue triangles control for % urban as well.

controls, and suggests that a percentage point increase in overhead efficiency is associated with a 0.2 percentage point increase in the Democratic vote. Including the standard controls from Table 1 results in a somewhat smaller but statistically significant coefficient in (2). These results suggest that this metric of efficiency, which is correlated with urbanization, was also correlated with voting. Models (3) and (4) also control for urbanization. Doing so results

	Δ% Democrat 1932–1936			
	(1)	(2)	(3)	(4)
overhead efficiency	0.199***	0.119***	0.076**	0.104**
	(0.036)	(0.040)	(0.032)	(0.039)
% urban			0.110***	0.033***
			(0.008)	(0.010)
Controls		x		x
DV mean	-0.011	-0.011	-0.011	-0.011
N	1756	1754	1754	1754
R^2	0.426	0.581	0.531	0.585

This table shows the results of regressions of the county-level change in the Democrats' share of the two-party vote between 1932 and 1936 against the overhead efficiency, calculated as total spending on government provision exclusive of central administrative costs divided by total spending inclusive of administrative costs. All models include state fixed effects. Even-numbered models also control for the shares employed in agriculture and manufacturing, the white and immigrant shares of the population, and union potential. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 2: Effects of public sector efficiency on realignment towards the Democrats after 1932

in very similar point estimates.[17] In Models (3) and (4), the coefficient on % urban is positive and similar to the estimated coefficients in Table 1. This result is to be expected given that the overhead metric only picks up a specific form of efficiency correlated with urbanization, and one would expect there to be many other ways that urbanization affects efficiency and thus voting.

Figure 5 shows the relationship between overhead efficiency and voting over time, given county and state-by-year fixed effects and the controls from model (4) interacted with year indicators. The effect of this measure of efficiency in 1932 persists into the 1950s, and there is not clear evidence of high-efficiency locations following differential trends prior to the 1932–1936 realignment. The one exception is the 1912 election, for which the dependent variable excludes support for Theodore Roosevelt's Progressive candidacy, and compares the economically-conservative Republican Taft against the more progressive Wilson. This particular comparison arguably captures the starkest left-right division on economic issues

[17]These results are also robust to controlling for a range of other covariates (Table A-9), and to not winsorizing overhead efficiency (Table A-10). County-level overhead efficiency is not driven by a taste for higher spending: it is uncorrelated with pre-New Deal spending after adjusting for population (Table A-12).

prior to the New Deal.

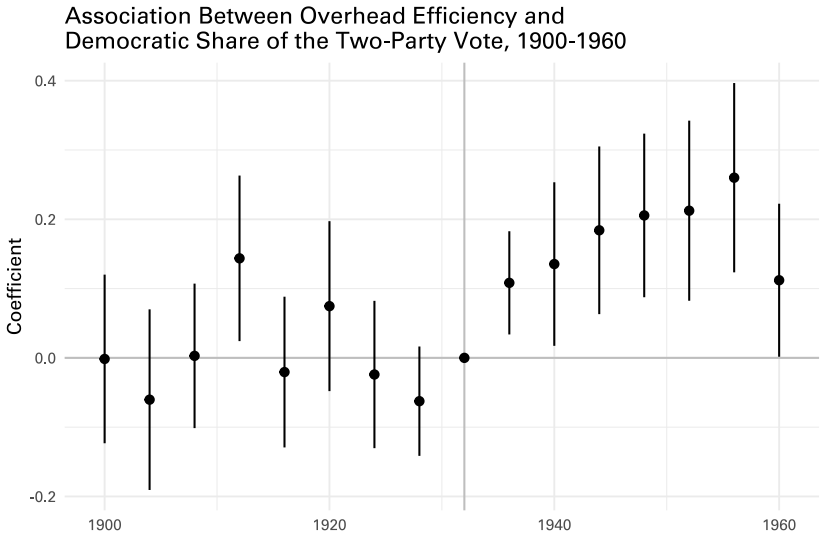

Figure 5: Overhead efficiency is associated with shifts towards the Democrats after 1932, but not before

This figure shows point estimates and 95% confidence intervals from a regression of the Democrats' share of the two-party vote on overhead efficiency interacted with year indicators, controlling for county and state-by-year fixed effects and controls from Table 2 model (4) interacted with year indicators.

Appendix B provides evidence for a number of other ways in which agglomeration was associated with more efficient government. Larger towns and more urban counties spent a smaller share of their pre-New Deal budgets on highway maintenance, a semi-rivalrous public good subject to clear economies of scale (Tables A-1 and A-2, Figure A-1). Such places experienced greater provision, as measured by police and firefighting employment per capita, relative to spending on those services. These two efficiency measures were also associated with the shift towards the Democrats post-1932 (Table A-11). The implementation of the New Deal was also more efficient in more urban environments. The share of those eligible enrolled in Social Security in 1940, and the probability that a town contained non-zero employment in New Deal Emergency Work schemes, increased sharply with town size (Figure A-2). The relationship between town size and the rate of participation in Emergency Work schemes followed an inverse u-shaped relationship, consistent with the presence of spatial barriers to access and local congestion costs. The cost of administering the Social Security

Employment Service, as measured by the ratio of administrative expenses to unemployment compensation actually paid, was lower in more urban states (Figure A-3). In Michigan, the ratio of administrative costs to relief disbursements was lower in counties with higher rates of agglomeration (Table A-5).

4 INDIVIDUAL-LEVEL EVIDENCE FOR THE NEW DEAL REALIGNMENT

I use polling data from 1936 to 1938 to test the individual-level predictions of the model, and probe whether that model can explain the 1932–1936 realignment. These polls were conducted by George Gallup's American Institute of Public Opinion through a quota-controlled sampling method that explicitly underweighted non-voters. I use the survey weights developed by Berinsky et al. (2011) to approximate the overall US population average.

The direct prediction from Equation (2) is that voters in more rural locations should prefer less government spending and a lower tax rate. I therefore examine questions about whether voters want the government to increase or decrease spending, and make use of a question from a 1937 survey which asked respondents to specify how much income tax a hypothetical man with a given income ought to pay. Figure 6 shows coefficients and confidence intervals from regressions of agreement with survey questions on urban status, with and without controls for race, gender, occupation, and region (Table A-13 presents the same results). Consistent with the theory, urban voters were more supportive of government spending and preferred higher tax rates.

Urban support for higher taxes and spending was accompanied by closer political alignment with the New Deal and Roosevelt's policies. Establishing that urban voters did not just shift towards the Democrats, but also supported Roosevelt's policies helps link the theoretical model to the results on agglomeration and voting. The most informative question on this topic is one that asked respondents whether they would prefer a New Dealer or Conservative Democrat if Roosevelt did not run for President in 1940. A vote for Roosevelt in 1936 was not necessarily an endorsement of the New Deal, but the result that urban respondents identified

more with the New Deal wing of the Democrats suggests the 1932–1936 shift in urban-rural voting was at least in part related to these policies. Questions on liberal and conservative policies give a similar impression. It is important to note that despite the apparent urban-rural division in support for liberal policies and a proposed liberal party, there is no relationship between rural and urban status and identification as a liberal or conservative, suggesting that policy preferences and not perhaps slower-changing political identity accounts for the emerging urban-rural divide.

Last, I examine support for specific New Deal policies. The theory suggests that urban respondents' support for spending should be concentrated on policies that have to be implemented on the ground, for which the greater efficiency of government in higher-agglomeration locations is important. Consistent with that intuition, rural respondents were less supportive of unemployment relief, which mainly took the form of job creation on public works projects, and the Tennessee Valley Authority. One would not expect to observe differences between urban and rural voters' attitudes to policies without a place-based component. I find no difference between urban and rural respondents in support for old age pensions and farm benefits, and find that rural respondents, perhaps unsurprisingly, were more supportive of the Agricultural Adjustment Act, which paid farmers to reduce crop production. These results suggest that it was not differences in attitudes to government spending and redistribution in some abstract sense that account for the realignment. An interesting exception concerns support for Social Security, which was more concentrated among urban respondents, even when controlling for occupation. The 1935 Social Security Act included provisions for the unemployed, dependent children, the blind, and public health, as well as for the elderly. Enrollment in Social Security was considerably higher in more urban areas, even accounting for the program's exclusion of agricultural workers (Figure A-2), suggesting that rural people faced barriers to accessing the program. According to the Social Security Board, the decision to exclude agricultural workers was due to "administrative difficulties" as "a large number of individual workers are employed in small establishments scattered over a wide area, frequently

at some distance from any city or town" (Committee on Economic Security, 1937, 208).[18]

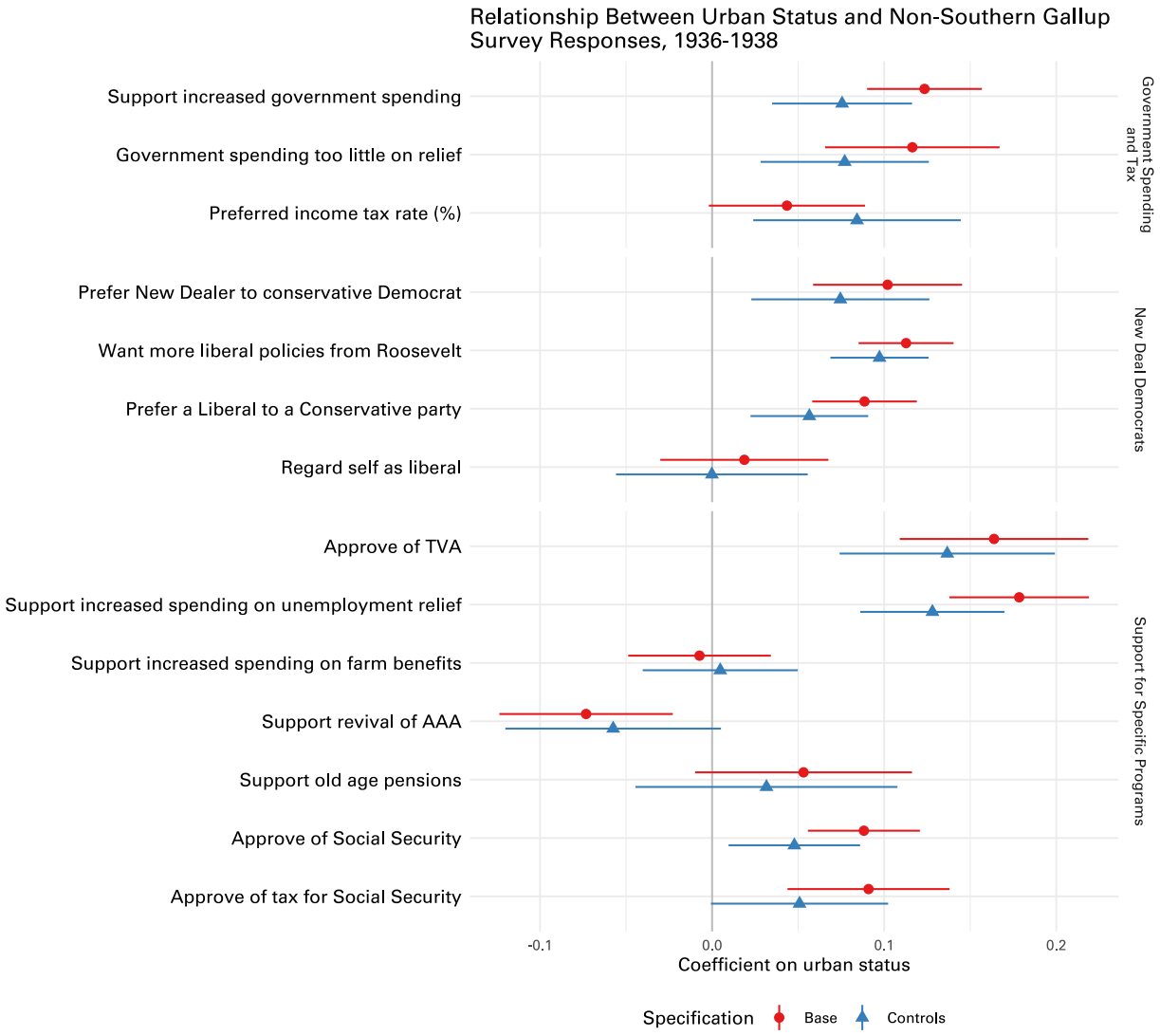

Figure 6: Urban survey respondents were more supportive of the New Deal
This figure shows coefficients and 95% confidence intervals from individual-level regressions of agreement with the survey question (coded 0–1) against urban status. Blue triangles control for race, gender, and occupation.

5 BRITAIN

While the analysis thus far has focused entirely on the US, the theory is not specific to that case. I examine the evolution of the rural-urban divide in voting in the UK over the

[18] On the debate about exclusions from social security, see DeWitt (2010).

late-19th and early-20th centuries. I assemble a constituency-level dataset using electoral data from Eggers and Spirling (2014) and microdata from the 1911 Census of England and Wales (Schürer and Higgs, 2014), which I link to a parish-level GIS from the Great Britain Historical GIS and spatially merge to 1885–1910 and 1918–1935 constituencies. I use the census data to calculate the shares of the population employed in agriculture, working class occupations, and mining and manufacturing, which had high rates of unionization.

The emergence of redistribution as the dominant issue in politics was more gradual in the UK than in the US. In 1906, the Liberal Party—which had previously been committed to classical liberalism—won a landslide election and formed a government that implemented a series of social welfare programs such as old age pensions and workers' compensation. If redistribution influenced the urban-rural divide, one would expect urban areas to shift towards the Liberals and away from the Conservatives in 1906. In the 1920s, the Labour Party, which was explicitly committed to redistribution, replaced the Liberals as the dominant party of the left. In 1918, the Conservative and Liberal parties, though internally-divided, campaigned to maintain the First World War coalition headed by the Liberal leader Lloyd George. This coalition was relatively centrist on economic issues and argued for welfarist reforms. In 1922 the Conservative Party withdrew from the coalition and won the resulting election on a manifesto promising fiscal retrenchment.

Table 3 examines the emergence of the urban-rural divide in the UK in these two instances. In models (1) and (2) the dependent variable is the change in the Conservative vote in constituencies in England and Wales, which picks up the main left-right divide, between 1900 and 1906, in (3) and (4) the change between 1918 and 1922. In both cases, a widening division between the parties on redistribution was associated with a shift towards the Conservatives by geographically-larger constituencies. As constituencies were apportioned to have similar populations, log constituency area essentially captures the log inverse of density, comparing compact urban constituencies to sparse rural ones. These associations hold up to controlling for agriculture, working class occupations, and sectors of the economy with high

	1900–1906		1918–1922	
	(1)	(2)	(3)	(4)
log area	1.073***	1.027**	1.502***	1.813*
	(0.220)	(0.438)	(0.480)	(0.989)
Controls		x		x
DV mean	-11.66	-11.66	0.248	0.248
N	292	292	414	414
R^2	0.054	0.096	0.024	0.054

This table shows the results of constituency-level regressions of the change in the Conservative vote share between 1900 and 1906 and between 1918 and 1922 on the log of constituency area. Models (2) and (4) also control for the share employed in agriculture, the share employed in manufacturing and mining, and the share in working class occupations. Robust standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 3: Effects of constituency area on shifts in Conservative voting in the UK

rates of union activity, suggesting that part of the contribution was by urbanization and not factors that simply happened to be correlated with urbanization. Figure 7 shows coefficients from a regression of the Conservative vote against log area interacted with year indicators, with and without the other controls interacted with year indicators. The association between density and left-wing voting only emerged in the 1920s with the rise of Labour; if anything, in the 1890s the Conservatives performed better in denser areas.

Canada offers another case of the urban-rural divide accompanying welfare state expansion. Armstrong, Lucas and Taylor (2022) document that the familiar urban-rural divide, in which rural areas vote conservative, emerged in Canada in the 1960s. This was a period in which the center-left Liberal Party moved left in response to electoral threat from the New Democratic Party, and set up a number of new welfare programs, including a single-payer health system (Johnston, 2013).

6 CONCLUSION

This paper introduces a new mechanism influencing preferences over the size of government and a new explanation for the urban-rural divide. Greater efficiency in government provision

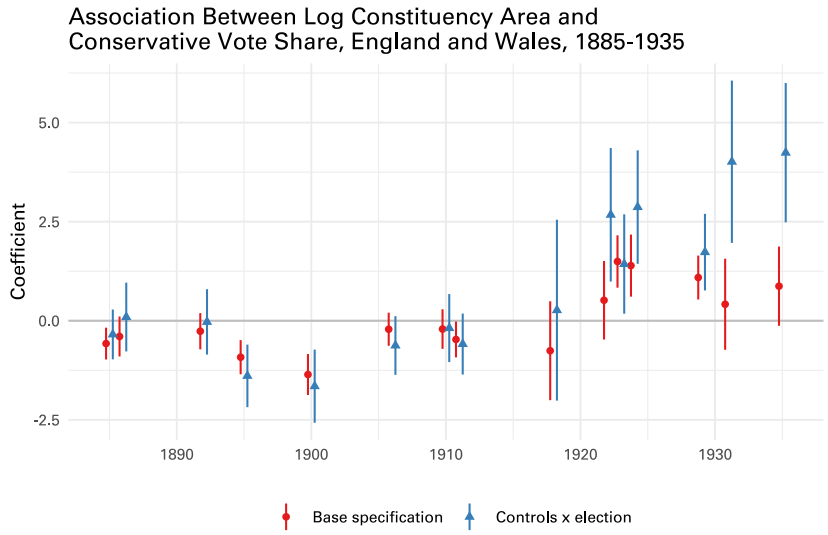

Figure 7: Negative association between log constituency area and voting for the Conservative Party in the UK with the rise of the welfare state and Labour

This figure shows coefficients and 95% confidence intervals from a regression of Conservative vote share on log constituency area interacted with election indicators. Red circles include election fixed effects, blue triangles add controls for employment in agriculture, manufacturing and mining, and working class occupations interacted with election indicators.

in cities due to economies of scale, nonrivalries, and access makes urban dwellers willing to accept higher taxes in return for more government spending. When redistribution became a salient issue dividing Republicans and Democrats between 1932 and 1936, more urban areas, and areas with more efficient local governments, shifted towards the Democrats, even when adjusting for the effects of other factors such as the presence of agriculture, manufacturing, union-intensive industries, and ethnic minorities. This result is corroborated by individual-level survey data, which suggests that urban voters in this period were more supportive of government expansion and the New Deal. The emergence of urban-rural electoral divides in other cases is consistent with this mechanism.

These results suggest a model of cleavage formation. As in Rogowski's (1989) account, political alignments emerge from the sharp distributional consequences of government policies. Once it became possible for states to tax and redistribute, the extent of redistribution had large consequences for all voters that influenced how they voted. This article draws attention to the importance of spatial variation in government efficiency for voters' preferences with

respect to redistribution. It thus provides an explanation for why political conflict plays out not just between classes, industries, ethnicities, and other interest groups, but also across space (Gimpel et al., 2020). It makes sense of why Rogowski's theory of political alignment characterizes political conflict in the 19th century well, but is inconsistent with patterns in the 20th century. Urban-rural and class cleavages are not alternative outcomes—as Rogowski's model suggests—but the joint result of redistributive politics. Beyond redistribution, the spatial distribution of the efficiency of different policy instruments, as well as that of interest groups, should influence the kinds of spatial divisions that emerge.

This analysis has implications for the literatures on redistribution and trade. Public agglomeration effects suggest an explanation for rising economic inequality. Whether local population change increases or decreases support for redistribution depends on whether it affects voters' incomes more than the efficiency of government provision. The spatial economics literature has tended to find weaker agglomeration effects in predominantly rural sectors like agriculture than in urban sectors like services (Kline and Moretti, 2014; Glaeser and Gottlieb, 2009). Thus as rural areas depopulate, one would expect them to shift right as public agglomeration effects dominate private ones. With respect to the politics of trade, scholarship following Ruggie (1982) links support for openness to the commitment of states to compensate those harmed by globalization. That it is harder for states to compensate rural voters suggests a reason why states tend to protect the agricultural sector (Davis, 2004), and why the response of—predominantly rural—areas harmed economically by trade with China has been to support protectionist rather than welfarist policies (Autor et al., 2020). It may be more efficient for governments to help rural voters with protectionism rather than compensation. This paper introduces the idea that agglomeration effects in public service provision matter for politics. Enumerating the ways they matter, and exploring the other interactions between policy instruments and space, will require many more.

Acemoglu, Daron and James Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. Cambridge: Cambridge University Press.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm and Nikolaus Wolf. 2015. "The Economics of Density: Evidence From the Berlin Wall." *Econometrica* 83(6):2127–2189.

Alesina, Alberto and Enrico Spolaore. 2003. *The Size of Nations*. Cambridge, MA: MIT Press.

Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114(4):1243–1284.

Allen, Treb and Dave Donaldson. 2020. Persistence and Path Dependence in the Spatial Economy. Working Paper 28059 National Bureau of Economic Research.

Alt, James and Torben Iversen. 2017. "Inequality, Labor Market Segmentation, and Preferences for Redistribution." *American Journal of Political Science* 61(1):21–36.

Anderson, William. 1936. *The Units of Government in the United States: An Enumeration and Analysis*. Chicago: Public Administration Service.

Armstrong, David A., Jack Lucas and Zack Taylor. 2022. "The Urban-Rural Divide in Canadian Federal Elections, 1896–2019." *Canadian Journal of Political Science/Revue canadienne de science politique* 55(1):84–106.

Autor, David, David Dorn, Gordon Hanson and Kaveh Majlesi. 2020. "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure." *American Economic Review* 110(10):3139–3183.

Bartolini, Stefano and Peter Mair. 1990. *Identity, Competition, and Electoral Availability:*

The Stabilisation of European Electorates 1885-1985. Cambridge: Cambridge University Press.

Benabou, Roland and Efe A. Ok. 2001. "Social Mobility and the Demand for Redistribution: The POUM Hypothesis." *Quarterly Journal of Economics* 116(2):447–487.

Berinsky, Adam J., Eleanor Neff Powell, Eric Schickler and Ian Brett Yohai. 2011. "Revisiting Public Opinion in the 1930s and 1940s." *PS: Political Science & Politics* 44(3):515–520.

Berkes, Enrico, Ezra Karger and Peter Nencka. 2022. "The census place project: A method for geolocating unstructured place names." *Explorations in Economic History* p. 101477.

Bleakley, Hoyt and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127(2):587–644.

Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge: Cambridge University Press.

Broz, J. Lawrence, Jeffry Frieden and Stephen Weymouth. 2021. "Populism in Place: The Economic Geography of the Globalization Backlash." *International Organization* 75(2):464–494.

Caramani, Daniele. 2004. *The Nationalization of Politics: The Formation of National Electorates and Party Systems in Western Europe*. Cambridge: Cambridge University Press.

Carmines, Edward G. and James A. Stimson. 1989. *Issue Evolution: Race and the Transformation of American Politics*. Princeton: Princeton University Press.

Carruthers, John I and Gudmundur F Ulfarsson. 2003. "Urban Sprawl and the Cost of Public Services." *Environment and Planning B: Planning and Design* 30(4):503–522.

Caughey, Devin, Michael C. Dougal and Eric Schickler. 2020. "Policy and Performance in the New Deal Realignment: Evidence from Old Data and New Methods." *Journal of Politics* 82(2):494–508.

Ciccone, Antonio and Robert E. Hall. 1996. "Productivity and the Density of Economic Activity." *The American Economic Review* 86(1):54–70.

Clubb, Jerome M., William H. Flanigan and Nancy H. Zingale. 2006. "Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972: Version 1.".

URL: *https://www.icpsr.umich.edu/web/ICPSR/studies/8611/versions/V1*

Collins, William J. and Gregory T. Niemesh. 2019. "Unions and the Great Compression of wage inequality in the US at mid-century: evidence from local labour markets." *Economic History Review* 72(2):691–715.

Committee on Economic Security. 1937. *Social Security in America*. Washington DC: Social Security Board.

Cramer, Katherine J. 2016. *The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker*. Chicago: University of Chicago Press.

Dasgupta, Aditya and Elena Ruiz Ramirez. 2020. Explaining Rural Conservatism: Political Consequences of Technological Change in the Great Plains. Technical report SocArXiv.

Davis, Christina L. 2004. "International Institutions and Issue Linkage: Building Support for Agricultural Trade Liberalization." *American Political Science Review* 98(1):153–169.

Derenoncourt, Ellora. 2022. "Can You Move to Opportunity? Evidence from the Great Migration." *American Economic Review* 112(2):369–408.

DeWitt, Larry. 2010. "The Decision to Exclude Agricultural and Domestic Workers from the 1935 Social Security Act." *Social Security Bulletin* 70(4):49–68.

Duranton, Gilles and Diego Puga. 2004. Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. In *Handbook of Regional and Urban Economics*, ed. J. Vernon Henderson and Jacques-François Thisse. Elsevier pp. 2063–2117.

Eggers, Andrew C. and Arthur Spirling. 2014. "Electoral Security as a Determinant of Legislator Activity, 1832–1918: New Data and Methods for Analyzing British Political Development." *Legislative Studies Quarterly* 39(4):593–620.

Enos, Ryan D. 2016. "What the Demolition of Public Housing Teaches Us about the Impact of Racial Threat on Political Behavior." *American Journal of Political Science* 60(1):123–142.

Federal Security Agency. 1941. *Fifth Annual Report of the Social Security Board 1940*. Washington: United States Government Printing Office.

Fetter, Daniel. 2017. "Local government and old-age support in the New Deal." *Explorations in Economic History* 66:1–20.

Fishback, Price V., Shawn Kantor and John Joseph Wallis. 2003. "Can the New Deal's three Rs be rehabilitated? A program-by-program, county-by-county analysis." *Explorations in Economic History* 40(3):278–307. Number: 3.

Fishback, Price V., William C. Horrace and Shawn Kantor. 2005. "Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales During the Great Depression." *Journal of Economic History* 65(1):36–71.

Ford, Robert and Will Jennings. 2020. "The Changing Cleavage Politics of Western Europe." *Annual Review of Political Science* 23(1):295–314.

Galbraith, J. K. 1934. California County Expenditures. Bulletin 582 University of California College of Agriculture, Agricultural Experimentation Station Berkeley, California: .

Gelman, Andrew, David Park, Boris Shor and Jeronimo Cortina. 2010. *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*. Princeton: Princeton University Press.

Gethin, Amory, Clara Martínez-Toledano and Thomas Piketty. 2022. "Brahmin Left Versus

Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948–2020." *Quarterly Journal of Economics* 137(1):1–48.

Gimpel, James G., Nathan Lovin, Bryant Moy and Andrew Reeves. 2020. "The Urban–Rural Gulf in American Political Behavior." *Political Behavior* 42(4):1343–1368.

Glaeser, Edward L. and Joshua D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47(4):983–1028.

Haber, William. 1935. *Cost of Administration in the Emergency Relief Program*. Lansing, MI: State Emergency Welfare Relief Commission.

Hirano, Shigeo and James M. Snyder. 2007. "The Decline of Third-Party Voting in the United States." *Journal of Politics* 69(1):1–16.

Hiscox, Michael J. 1999. "The Magic Bullet? The RTAA, Institutional Reform, and Trade Liberalization." *International Organization* 53(4):669–698.

Hiscox, Michael J. 2002. *International Trade and Political Conflict: Commerce, Coalitions, and Mobility*. Princeton: Princeton University Press.

Hooghe, Liesbet and Gary Marks. 2018. "Cleavage theory meets Europe's crises: Lipset, Rokkan, and the transnational cleavage." *Journal of European Public Policy* 25(1):109–135.

Huber, John D. and Piero Stanig. 2011. "Church-state separation and redistribution." *Journal of Public Economics* 95(7):828–836.

Inglehart, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton: Princeton University Press.

Johnston, Richard. 2013. The Party System, Elections, and Social Policy. In *Inequality and the Fading of Redistributive Politics*, ed. Keith G. Banting and John Myles. Vancouver: UBC Press.

Key, V. O. 1966. *The Responsible Electorate: Rationality in Presidential Voting, 1936–1960.* Harvard University Press.

Key, V.O. 1955. "A Theory of Critical Elections." *Journal of Politics* 17(1):3–18.

Kinder, Donald R. and D. Roderick Kiewiet. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science* 11(2):129–161.

Kline, Patrick and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics* 129(1):275–331.

Kriesi, Hanspeter. 1998. "The transformation of cleavage politics The 1997 Stein Rokkan lecture." *European Journal of Political Research* 33(2):165–185.

Lancaster, Lane W. 1937. *Government in Rural America.* New York: D. Van Nostrand Company, Inc.

Lee, Woojin and John E. Roemer. 2006. "Racism and redistribution in the United States: A solution to the problem of American exceptionalism." *Journal of Public Economics* 90(6):1027–1052.

Lehman, Pola, Simon Franzmann, Tobias Burst, Sven Regel, Felicia Riethmüller, Andrea Volkens, Bernhard Weßels and Lisa Zehnter. 2023. "Manifesto Project (MRG/CMP/MARPOR).".
URL: *https://doi.org/10.25522/manifesto.mpds.2023a*

Lindbeck, Assar and Jörgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice* 52(3):273–297.

Lipset, Seymour Martin and Stein Rokkan. 1967. Cleavage Structures, Party Systems and Voter Alignments: An Introduction. In *Party Systems and Voter Alignments: Cross-National Perspectives*, ed. Seymour Martin Lipset and Stein Rokkan. New York: Free Press pp. 1–64.

Lizzeri, Alessandro and Nicola Persico. 2004. "Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform'." *Quarterly Journal of Economics* 119(2):707–765.

Luebbert, Gregory M. 1991. *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe*. New York: Oxford University Press.

Mansfield, Edward D. and Diana C. Mutz. 2009. "Support for Free Trade: Self-Interest, Sociotropic Politics, and Out-Group Anxiety." *International Organization* 63(3):425–457.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler and Steven Ruggles. 2020. "IPUMS National Historical Geographic Information System: Version 15.0 [dataset].". **URL:** *http://doi.org/10.18128/D050.V15.0*

Manza, Jeff and Clem Brooks. 1999. *Social Cleavages and Political Change: Voter Alignments and U.S. Party Coalitions*. New York: Oxford University Press.

Meltzer, Allan H. and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5):914–927.

Moene, Karl Ove and Michael Wallerstein. 2001. "Inequality, Social Insurance, and Redistribution." *American Political Science Review* 95(4):859–874.

Moore, Barrington. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Norpoth, Helmut. 2019. "The American Voter in 1932: Evidence from a Confidential Survey." *PS: Political Science & Politics* 52(1):14–19.

Oates, Wallace E. 1972. *Fiscal Federalism*. New York: Harcourt Brace Jovanovich, Inc.

Ogorzalek, Thomas K. 2018. *The Cities on the Hill: How Urban Institutions Transform National Politics*. New York: Oxford University Press.

Olken, Benjamin A. 2006. "Corruption and the costs of redistribution: Micro evidence from Indonesia." *Journal of Public Economics* 90(4):853–870.

Paluck, Elizabeth Levy, Seth A. Green and Donald P. Green. 2019. "The contact hypothesis re-evaluated." *Behavioural Public Policy* 3(2):129–158.

Piketty, Thomas. 1995. "Social Mobility and Redistributive Politics." *The Quarterly Journal of Economics* 110(3):551–584.

Posner, Daniel N. 2004. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi." *American Political Science Review* 98(4):529–545.

Rickard, Stephanie J. 2020. "Economic Geography, Politics, and Policy." *Annual Review of Political Science* 23(1):187–202.

Rodden, Jonathan. 2019. *Why Cities Lose: The Deep Roots of the Urban-Rural Political Divide*. New York: Hachette.

Roemer, John E. 2006. *Political Competition: Theory and Applications*. Cambridge, MA; London: Harvard University Press.

Rogowski, Ronald. 1989. *Commerce and Coalitions: How Trade Affects Domestic Political Alignments*. Princeton: Princeton University Press.

Romer, Thomas. 1975. "Individual welfare, majority voting, and the properties of a linear income tax." *Journal of Public Economics* 4(2):163–185.

Rueda, David and Daniel Stegmueller. 2016. "The Externalities of Inequality: Fear of Crime and Preferences for Redistribution in Western Europe." *American Journal of Political Science* 60(2):472–489.

Ruggie, John Gerard. 1982. "International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order." *International Organization* 36(2):379–415.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler and Matthew Sobek. 2021. "IPUMS USA: Version 11.0 [dataset].".
URL: *https://doi.org/10.18128/D010.V11.0*

Sands, Melissa L. 2017. "Exposure to inequality affects support for redistribution." *Proceedings of the National Academy of Sciences* 114(4):663–668.

Scheve, Kenneth and David Stasavage. 2006. "Religion and Preferences for Social Insurance." *Quarterly Journal of Political Science* 1(3):255–286.

Schickler, Eric. 2016. *Racial Realignment: The Transformation of American Liberalism, 1932–1965*. Princeton: Princeton University Press.

Schlesinger, Arthur M. 1957. *The Age of Roosevelt: The Crisis of the Old Order, 1919-1933*. Boston: Houghton Mifflin Co.

Schürer, Kevin and Edward Higgs. 2014. "Integrated Census Microdata (I-CeM), 1851-1911.". SN: 7481.
URL: *http://doi.org/10.5255/UKDA-SN-7481-1*

Shayo, Moses. 2009. "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution." *American Political Science Review* 103(2):147–174.

Troy, Leo. 1957. *Distribution of Union Membership among the States, 1939 and 1953*. National Bureau of Economic Research, Inc.

US Department of Commerce. 1969. *Census of Governments, 1967, Vol.6: Topical Studies, No. 5: Historical Statistics of Governmental Finances and Employment*. Washington DC: US Government Printing Office.

Walker, Mabel. 1930. *Municipal Expenditures*. Baltimore: Johns Hopkins University Press.

Wallis, John Joseph. 1984. "The Birth of the Old Federalism: Financing the New Deal, 1932-1940." *Journal of Economic History* 44(1):139–159.

Wallis, John Joseph. 2000. "American Government Finance in the Long Run: 1790 to 1990." *Journal of Economic Perspectives* 14(1):61–82.

Online Appendix

LIST OF TABLES

LIST OF FIGURES

A A MODEL OF URBAN-RURAL PLATFORM DIVERGENCE AND CONVERGENCE WITH ENDOGENOUS COALITIONS

In the model in the main body of the paper, the convergence and divergence of the parties on policy is taken as given. The paper's core contribution does not rely on a particular reason for why the parties had converged on redistribution and subsequently diverged. The extension below sets out one way in which this pattern can be rationalized.

A.1 *Setup*

I impose a number of additional assumptions on the model developed in the main body of the paper to make the analysis of coalition formation tractable. There are only two types of locations, urban and rural. I normalize government productivity $\bar{A}n_l^\alpha = 1$ if l is rural, and $\bar{A}n_l^\alpha = A > 1$ if l is urban. I denote the share of the population living in urban locations by λ. The average value of ψ in rural areas is ψ_r and the average in urban areas is ψ_u. I assume all agents in a given location have the same income: urban agents have income $\bar{y}_u$, rural agents $\bar{y}_r$. Total average income is $\bar{y} = \lambda\bar{y}_u + (1-\lambda)\bar{y}_r$. I make the following assumptions on the relative magnitudes of these parameters:

1. The ratio of urban to rural public-sector productivity exceeds the ratio of urban to rural average income: $A > \frac{\bar{y}_u}{\bar{y}_r}$. This assumption ensures that urban voters benefit more than rural voters from higher taxes.

2. The average bias for party L is weakly higher among urban than rural voters: $\psi_u \geq \psi_r$. This assumption is the only exogenous difference between the parties—the parties' policy preferences are determined in equilibrium—and so is without loss of generality.

As in the main body of the paper, voters receive an individual uniform preference shock in addition to a place-specific preference shock. Vote share for party L in rural areas is

$$s_r(t^L, t^R) = \theta\bar{y}^\varphi\left((t^L)^\varphi - (t^R)^\varphi\right) - \theta(t^L - t^R)\bar{y}_r + \theta\psi_r + \frac{1}{2}$$

Vote share for party L in urban areas is

$$s_u(t^L, t^R) = \theta A\bar{y}^\varphi\left((t^L)^\varphi - (t^R)^\varphi\right) - \theta(t^L - t^R)\bar{y}_u + \theta\psi_u + \frac{1}{2}$$

Parties choose the policy that, if implemented, maximizes the utility of their median voter. In the language of Roemer (2006), both parties are run by "militants." A party that wins more votes from urban areas will choose a policy that maximizes the utility of urban voters; one that wins more votes from rural areas will choose a policy that maximizes the utility of rural voters. This assumption can be justified on the grounds that parties cannot credibly commit before an election to implement policies, but once in office are accountable to their memberships.

The policy that maximizes utility for urban voters is

$$t_u^* = \arg\max_t A\left(t\bar{y}\right)^{\varphi} + (1-t)\bar{y}_u = \left(\frac{\varphi A \bar{y}^{\varphi}}{\bar{y}_u}\right)^{\frac{1}{1-\varphi}}$$

The policy that maximizes utility for rural voters is

$$t_r^* = \left(\frac{\varphi \bar{y}^{\varphi}}{\bar{y}_r}\right)^{\frac{1}{1-\varphi}}$$

We can write

$$t_u^* = t_r^* \left(\frac{A\bar{y}_r}{\bar{y}_u}\right)^{\frac{1}{1-\varphi}}$$

The ratio of the preferred urban tax rate relative to the preferred rural tax rate is increasing in A, the relative productivity of public goods provision in urban settings, and decreasing in $\frac{\bar{y}_u}{\bar{y}_r}$, the ratio of urban to rural average income. Because by assumption $A > \frac{\bar{y}_u}{\bar{y}_r}$, we see $t_u^* > t_r^*$. If the efficiency of providing the public good in urban areas exceeds the difference in wealth between urban and rural areas, urban voters will demand higher tax rates than rural voters.

A.2 *Equilibrium*

A political equilibrium is a pair of policies $\{t^L, t^R\}$ and accompanying vote shares $\{s_u(t^L, t^R), s_r(t^L, t^R)\}$ such that each party's policy maximizes the utility of the median voter within its coalition, and each voter chooses the party that maximizes her payoff subject to preference shocks.

In equilibrium, the parties can converge on the policies preferred by rural or urban voters, or diverge. For intuition for why convergence can occur in equilibrium, suppose that voters' have no non-economic place-specific preference for either party, that is, $\psi_r = \psi_u = 0$. Under that assumption, if the parties choose the same policy, their coalitions will be identical. Divergence can be an equilibrium if the party that proposes the higher tax rate wins most of its votes from urban voters, and the party that proposes the lower tax rate wins most of its votes from rural voters.

Proposition 1. *If* $\lambda \leq \bar{\lambda}_r := \frac{1+2\theta\psi_r}{2(1+\theta(\psi r+\psi_u))}$, $\{t_r^*, t_r^*\}$ *is a political equilibrium.*

For there to be a political equilibrium in which both parties maximize the welfare of rural voters, the median voter in both parties must be rural. From the assumption that $\psi_r \leq \psi_u$, if both parties choose the same policy, party R will win more votes from rural voters relative to urban voters than party L. If L's median voter is rural, so is R's. The fraction of the total

vote L wins from rural voters is given by $(1-\lambda)s_r(t_r^*, t_r^*)$, and the fraction of the total vote L wins from urban voters is given by $\lambda s_u(t_r^*, t_r^*)$. Combining these two expressions into an inequality and rearranging gives

$$\frac{\lambda}{1-\lambda} \leq \frac{s_r(t_r^*, t_r^*)}{s_u(t_r^*, t_r^*)}$$

Substituting in $s_u(t_r^*, t_r^*) = \theta\psi_u + \frac{1}{2}$ and $s_r(t_r^*, t_r^*) = \theta\psi_r + \frac{1}{2}$ gives

$$\lambda \leq \frac{1 + 2\theta\psi_r}{2(1 + \theta(\psi r + \psi_u))}$$

Proposition 2. *If $\lambda \geq \underline{\lambda}_u := \frac{1-2\theta\psi_r}{2(1-\theta(\psi r+\psi_u))}$, $\{t_u^*, t_u^*\}$ is a political equilibrium.*

The logic is similar to the previous proposition. The median voter in both parties must be urban. The relevant inequality that ensures that R wins more votes from urban than rural voters is

$$\frac{\lambda}{1-\lambda} \geq \frac{1 - s_r(t_u^*, t_u^*)}{1 - s_u(t_u^*, t_u^*)}$$

As $1 - s_r(t_u^*, t_u^*) = -\theta\psi_r + \frac{1}{2}$ and $1 - s_u(t_u^*, t_u^*) = -\theta\psi_u + \frac{1}{2}$, the relevant bound on λ is

$$\lambda \geq \frac{1 - 2\theta\psi_r}{2(1 - \theta(\psi r + \psi_u))}$$

Remark 1. *$\underline{\lambda}_u \geq \bar{\lambda}_r$, and if $\psi_u > \psi_r$ the inequality is strict.*

This remark follows directly from comparing the expressions for $\underline{\lambda}_u$ and $\bar{\lambda}_r$. Substantively, this result means that, except in the knife-edge case with no place-specific differences in bias between the parties, there cannot be multiple equilibria for a given value of λ in which both parties coordinate on either policy platform. If the bias terms differ between the parties, then there is a range of values of λ under which policy convergence is not an equilibrium. Under policy convergence, the difference in vote shares between the two parties are entirely due to the bias terms ψ, and so whether the parties converge on the policy preferred by rural or urban voters does not change vote shares. Thus, if the parties are converged and we increase λ slightly above the highest value for which both parties' majorities are rural, the party with the more positive bias among urban voters begins to win a majority of its votes from urban voters before the party with the less positive bias does.

Proposition 3. *$\exists \underline{\lambda}_d, \bar{\lambda}_d$ such that if $\lambda \in [\underline{\lambda}_d, \bar{\lambda}_d]$, $\{t_u^*, t_r^*\}$ is an equilibrium.*

For an equilibrium in which the parties diverge on policy, a majority of votes for L must come from urban voters and a majority of votes for R must come from rural voters. The first condition implies the inequality $\lambda s_u(t_u^*, t_r^*) \geq (1-\lambda)s_r(t_u^*, t_r^*)$, the second, $\lambda(1 - s_u(t_u^*, t_r^*)) \leq (1-\lambda)(1 - s_r(t_u^*, t_r^*))$. Define $\underline{\lambda}_d$ (d for divergence) as the lowest value of λ that satisfies the first inequality, and $\bar{\lambda}_d$ as the largest value of λ that satisfies the second. Rearranging and combining gives the following chain of inequalities:

$$\frac{s_r(t_u^*, t_r^*)}{s_u(t_u^*, t_r^*)} = \frac{\underline{\lambda}_d}{1 - \underline{\lambda}_d} \leq \frac{\lambda}{1-\lambda} \leq \frac{\bar{\lambda}_d}{1 - \bar{\lambda}_d} = \frac{1 - s_r(t_u^*, t_r^*)}{1 - s_u(t_u^*, t_r^*)}$$

To show that there are values of λ for which the divergent equilibrium exists, we must show that $\underline{\lambda}_d \leq \bar{\lambda}_d$.[19]

Note that $s_r(t_u^*, t_r^*) < s_r(t_r^*, t_r^*)$ and $s_u(t_u^*, t_r^*) > s_u(t_r^*, t_r^*)$. t_u^* is the policy that maximizes the welfare of urban voters, t_r^* is the policy that maximizes the welfare of rural voters, and the vote share for a given party with a given group of voters is increasing in the welfare of those voters under the policy the party proposes. If R chooses the policy t_r^*, and L switches from proposing t_r^* to t_u^*, L must increase its share of the vote among urban voters and decrease its share of the vote among rural voters. By the same logic, $1 - s_r(t_u^*, t_r^*) > 1 - s_r(t_u^*, t_u^*)$ and $1 - s_u(t_u^*, t_r^*) < 1 - s_u(t_u^*, t_u^*)$. Combining these inequalities with the identities of $\bar{\lambda}_r$ and $\underline{\lambda}_u$ gives the following:

$$\frac{s_r(t_u^*, t_r^*)}{s_u(t_u^*, t_r^*)} < \frac{s_r(t_r^*, t_r^*)}{s_u(t_r^*, t_r^*)} = \frac{\bar{\lambda}_r}{1 - \bar{\lambda}_r} \text{ and } \frac{1 - s_r(t_u^*, t_r^*)}{1 - s_u(t_u^*, t_r^*)} > \frac{1 - s_r(t_u^*, t_u^*)}{1 - s_u(t_u^*, t_u^*)} = \frac{\underline{\lambda}_u}{1 - \underline{\lambda}_u}$$

The left-hand side expression implies that $\underline{\lambda}_d < \bar{\lambda}_r$, the right-hand side that $\bar{\lambda}_d > \underline{\lambda}_u$. From Remark 1, it follows that $\underline{\lambda}_d < \bar{\lambda}_d$.

Remark 2. *For $\lambda \in [\underline{\lambda}_d, \bar{\lambda}_r]$ and $\lambda \in [\underline{\lambda}_u, \bar{\lambda}_d]$, both policy divergence and convergence are equilibria.*

This result follows immediately from the three propositions. Substantively this result means that at intermediate levels of urbanization one could observe polarization without any fundamental change in urbanization or other model parameters. The logic for why both divergent and convergent equilibria are possible at certain levels of urbanization is that policy divergence accentuates urban-rural divides. If a large enough majority of voters are rural, when both parties choose policies that benefit rural voters both win most of their votes from rural voters, creating an internal majority in favor of pro-rural policies in both parties. However, if the parties polarize, the party choosing policies that benefit urban voters more will bring more urban voters into its coalition, and lose rural voters, and so, if the rural population share is not too large, can create an internal majority in favor of pro-urban policies. A shock to a party's vote share among urban voters, for instance from implementing a policy that benefits urban voters more, can lead to durable policy divergence.

[19]Note that one can construct a similar chain of inequalities to specify the range of λ values under which $\{t_r^*, t_u^*\}$ is an equilibrium, that is, under which the party for which urban voters have a stronger taste proposes the policy preferred by rural voters, and wins most of its votes from rural voters, and the party for which rural voters have a stronger taste proposes the policy preferred by urban voters, and wins most of its votes from urban voters. Those conditions require $(1-\lambda)s_r(t_r^*, t_u^*) > \lambda s_u(t_r^*, t_u^*)$ and $\lambda(1 - s_u(t_r^*, t_u^*)) > (1-\lambda)(1 - s_r(t_r^*, t_u^*))$. There are values of λ for which such an equilibrium exists if the magnitude of urban-rural differences in non-economic preferences $\psi_u - \psi_r$ is small in relation to that of the urban-rural difference in policy preferences. More formally, the smallest λ that satisfies the first inequality is smaller than largest λ that satisfies the second if $s_r\left(t_r^*, t_u^*\right) > s_u\left(t_r^*, t_u^*\right)$, which gives the condition

$$\bar{y}^{\varphi}(A-1)\left((t_u^*)^{\varphi} - (t_r^*)^{\varphi}\right) - (\bar{y}_u - \bar{y}_r)(t_u^* - t_r^*) > \psi_u - \psi_r$$

Note also that the range of λ values under which $\{t_r^*, t_u^*\}$ is an equilibrium is smaller than the equivalent for $\{t_u^*, t_r^*\}$. The largest λ for $\{t_r^*, t_u^*\}$ must satisfy $\frac{\lambda}{1-\lambda} < \frac{s_r(t_r^*, t_u^*)}{s_u(t_r^*, t_u^*)}$, which is less than $\frac{1-s_r(t_u^*, t_r^*)}{1-s_u(t_u^*, t_r^*)}$, the equivalent condition for $\{t_u^*, t_r^*\}$. This comparison follows from the identities $s_r(t_r^*, t_u^*) = 1 - s_r(t_u^*, t_r^*) + 2\theta\psi_r$, $s_u(t_r^*, t_u^*) = 1 - s_u(t_u^*, t_r^*) + 2\theta\psi_u$, $\psi_u \geq \psi_r$, and $s_u(t_u^*, t_r^*) \geq s_r(t_u^*, t_r^*)$. One can use similar logic to show that the lowest value of λ for which $\{t_r^*, t_u^*\}$ is an equilibrium is greater than the lowest value of λ for $\{t_u^*, t_r^*\}$.

B ADDITIONAL EVIDENCE OF AGGLOMERATION EFFECTS IN GOVERNMENT PROVISION

This section presents additional evidence of the relationship between agglomeration and the efficiency of government provision before and during the New Deal. Counties with higher levels of agglomeration had more efficient local governments prior to the New Deal, a factor that is important given that many New Deal programs were in part implemented through local governments (Fetter, 2017). The provision of the New Deal was also less efficient in smaller localities.

In addition to the central administrative cost metric used in Section 3, I use the *Financial Statistics of State and Local Government* to study two other ways agglomeration affects government efficiency. First, I examine the share of the government budget going on highway maintenance. This outcome should pick up agglomeration effects related to access costs and nonrivalry. One would expect that in all locations the government spent money to provide the basic service of connecting residents to one another. Population sparsity should make doing so costly, in part because the number of people using each stretch of road would be fairly low. Figure A-1 shows the strong negative relationship at the town-level between population and the share of the operations budget going on highway maintenance. Columns (3) and (4) of Table A-1 show that larger towns spent a smaller share of their budgets on highway maintenance.

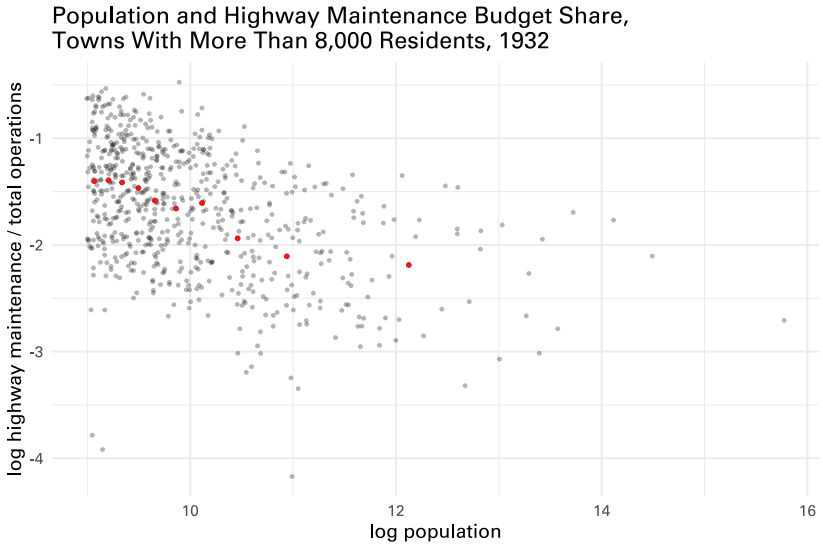

Figure A-1: Relationship between town population and the share of the government operations budget allocated to highway maintenance, non-southern towns with more than 8,000 residents in 1932 reporting government spending by category

Second, I examine the relationship between agglomeration and service provision conditional on spending. Substituting the agglomerative productivity expression into (1) and taking logarithms gives the following equation:

$$\ln G_l = \bar{A} + \alpha \ln n_l + \varphi \ln g_l$$

	overhead efficiency		highways		police & fire / pop	
	(1)	(2)	(3)	(4)	(5)	(6)
log pop	0.050***	0.044***	−0.283***	−0.267***	0.086***	0.078***
	(0.006)	(0.005)	(0.024)	(0.018)	(0.017)	(0.020)
log police & fire spend / pop					0.469***	0.512***
					(0.033)	(0.048)
State FE		x		x		x
DV mean	-0.235	-0.235	-1.676	-1.676	-6.706	-6.706
N	677	677	677	677	583	583
R^2	0.106	0.475	0.205	0.593	0.453	0.553

This table shows the results of regressions of government effciency against log town population, for non-Southern cities with more than 8,000 residents in 1932, for which spending by category is reported. In models (1) and (2) the dependent variable is log overhead efficiency, calculated as spending on government provision exclusive of central administrative costs and highway maintenance divided by total government operating spending net of highway maintenance, winsorized. In (3) and (4) the log ratio of highway maintenance spending to total government operating spending, in (5) and (6) the log number of police and firefighters divided by population. Even-numbered models add state fixed effects. Robust standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-1: Effects of town size on public sector efficiency, 1932

Conditional on log spending per capita (g_l), the log of town population (n_l) should be positively correlated with log provision per capita (G_l), and the coefficient on log population can be interpreted as α, the agglomeration elasticity. Models (5) and (6) of Table A-1 estimate this equation for the case of police and fire protection spending. This is a case in which service provision can arguably be measured by the number of police officers and firemen, which I measure at the town level using the 1930 Census microdata and the Census Place Project. Model (6) suggests an agglomeration elasticity of $\alpha = 0.078$. Table A-2 shows that this relationship holds at the county level, regressing the log of police and firemen per capita on the log of the agglomeration metric.[20] Table A-3 reports the results of a number of similar regressions linking various measures of service provision, such as school attendance, to agglomeration and spending. With the exceptions of the number of teachers and sanitation workers per capita, agglomeration is positively associated with better service provision across the board, controlling for category-specific spending. These results, especially as they pertain to outcomes like the employment of police, firefighters, and librarians, are somewhat surprising, in that higher nominal wages in cities should make hiring government employees more expensive. They suggest that economies of scale in administrative costs matter for how spending within a given category actually benefits citizens, as well as for the allocation of spending across categories. Counties with higher levels of agglomeration received more services for a given level of spending.

Residents of smaller towns and villages had more difficulty in accessing New Deal programs. The theoretical microfoundation for access costs in Section 2 suggests that as population density increases, the share of the population close enough to access government services increases, though access to that service might become more congested. Figure A-2 shows

[20]It also shows that the other results in Table A-1 hold at the county level.

	overhead efficiency		highways		police & fire / pop	
	(1)	(2)	(3)	(4)	(5)	(6)
log agglomeration	0.179***	0.214***	−1.773***	−2.087***	4.678***	2.151***
	(0.047)	(0.058)	(0.232)	(0.288)	(0.438)	(0.371)
log police & fire spend / pop					0.215***	0.195***
					(0.060)	(0.055)
Controls		x		x		x
DV mean	-0.189	-0.189	-2.057	-2.057	-7.374	-7.374
N	1751	1751	1742	1742	1579	1579
R^2	0.293	0.315	0.545	0.548	0.518	0.562

This table shows the results of county-level regressions of government spending and performance against log agglomeration. In models (1) and (2) the dependent variable is the log of overhead efficiency, calculated as government provision exclusive of central administrative costs and highway maintenance divided by total government operating spending net of highway maintenance, in (3) and (4), the log share of government operating spending allocated to highways, and in (5) and (6) the log number of police and firefighters divided by population. All models include state fixed effects, models (2), (4), and (6) control for the shares in agriculture and manufacturing, the immigrant and white population shares, and union potential. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-2: Relationship between agglomeration and government efficiency, for non-Southern counties in 1932

binned scatterplots at the place level in 1940 of the share of the unemployed population employed on emergency works programs, the probability that the place has at least one person employed in such a program, and the share of eligible residents enrolled in Social Security. There is a clear positive relationship between place size and Social Security enrollment (note that this variable excludes agricultural and domestic workers who were ineligible from the denominator). Employment in emergency works programs follows an inverse u-shaped relationship with place size, consistent with the predicted tradeoff between geographic access and congestion. At low levels of density, a large share of the population is not close enough to access geographically-limited services at all. At intermediate levels, the share of the population that is close enough to access is higher, though provision is relatively inefficient because there are relatively few people using each service location. At higher levels, the share accessing the service might decrease because of increased congestion. The share of locations with at least one person employed in emergency work is strongly increasing with place size. Thus in the smallest places many people may have been unaware of emergency works programs actually helping people like them, which would not have been the case in larger cities.

Economies of scale in administrative costs also appear to have applied to the administration of Social Security. Figure A-3 plots the ratio of administrative expenses of the Social Security Employment Service relative to unemployment compensation actually paid by the service against state-level urbanization, using data digitized from the 1940 Social Security Board annual report. In more urban states the administrative cost per dollar disbursed was lower.

In Michigan, for which Haber (1935) reports county level statistics on the administrative cost and total disbursements of the state's emergency relief program, administrative costs relative to disbursements were lower in counties with greater agglomeration (Table A-5).

	teachers	school enrol.	police	firemen	sanitation	inf. mortality	utilities	librarians	paved roads
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log agglomeration	−0.915***	0.126***	2.677***	5.578***	−0.799	−0.340**	3.320***	1.194***	5.947***
	(0.074)	(0.020)	(0.412)	(0.501)	(0.754)	(0.128)	(0.411)	(0.274)	(0.986)
log ed spend / pop	0.051								
	(0.037)								
log ed spend / child		0.019*							
		(0.011)							
log police & fire spend / pop			0.147**	0.353***					
			(0.060)	(0.089)					
log pub. health spend / pop					0.157				
					(0.102)				
log pub. health spend / births						0.034**			
						(0.017)			
log utilities spend / pop							0.072***		
							(0.018)		
log library spend / pop								0.217***	
								(0.039)	
log highway spend / farms									0.146**
									(0.065)
DV mean	-4.512	-0.508	-7.463	-8.698	-9.195	-2.871	-6.351	-8.452	-3.655
N	1756	1756	1756	1756	1756	1756	1756	1756	1754
R^2	0.600	0.492	0.393	0.476	0.245	0.214	0.416	0.402	0.731

This table shows the results of county-level regressions of various logged measures of service provision on the logarithm of agglomeration, controlling for sector-specific spending in 1932. In model (1) the dependent variable is the log number of teachers in 1930 per capita, in (2), the log share of children in school, in (3) the log number of police officers per capita, in (4) the log number of fire-fighters per capita, in (5) the log number of sanitation workers per capita, in (6) the log number of infant deaths over births in 1932, in (7) the log number of workers in public utilities per capita, in (8) the log number of librarians per capita, in (9) the log share of farms accessible by paved road. All models include state fixed effects. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-3: Relationship between agglomeration and service provision, controlling for spending, for counties in 1932

	log infant mortality			log school attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
overhead efficiency	-1.133^{***}	-0.868^{***}	-0.805^{***}	0.155^{***}	0.096^{***}	0.124^{***}
	(0.323)	(0.256)	(0.222)	(0.040)	(0.032)	(0.032)
log spending / pop			0.080			0.036^{***}
			(0.070)			(0.009)
Controls		x	x		x	x
DV mean	-2.866	-2.866	-2.866	-0.509	-0.509	-0.509
N	1760	1758	1758	1758	1758	1758
R^2	0.216	0.278	0.280	0.449	0.573	0.588

This table shows the results of county-level regressions of log infant mortality and the log share of children attending school against overhead efficiency. All models include state fixed effects, models (2), (3), (5), and (6) control for the shares in agriculture and manufacturing, the immigrant and white population shares, and union potential, (3) and (6) also control for log spending per capita. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-4: Relationship between overhead efficiency, infant mortality and school attendance, for non-Southern counties in 1932

	Administrative costs / total relief spending					
	(1)	(2)	(3)	(4)	(5)	(6)
agglomeration	-0.047^{***}	-0.043^{*}				
	(0.015)	(0.022)				
overhead efficiency			-0.184^{***}	-0.171^{***}	-0.155^{***}	-0.166^{***}
			(0.055)	(0.051)	(0.055)	(0.052)
% urban					-0.023^{**}	-0.014
					(0.009)	(0.012)
Controls		x		x		x
DV mean	0.081	0.081	0.081	0.081	0.081	0.081
N	82	82	82	82	82	82
R^2	0.141	0.241	0.142	0.312	0.203	0.321

This table shows the results of regressions of the county-level ratio of emergency relief administrative costs to total emergency relief spending, in Michigan, December 1934–February 1935, against agglomeration, overhead efficiency and urbanization. All models include an intercept. Even-numbered models also control for the shares employed in agriculture and manufacturing, the white and immigrant shares of the population, and union potential. Robust standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-5: Relationship between agglomeration, overhead efficiency, and the cost of emergency relief administration in Michigan

Figure A-2: Relationship between place size and New Deal participation

C ADDITIONAL TABLES AND FIGURES

Table A-6: County-level summary statistics

Variable	N	Mean	SD	q5	q95
% urban, 1930	1,761	0.253	0.266	0.000	0.763
agglomeration, 1930	1,758	1.834	0.211	1.584	2.232
agglomeration (winsorized), 1930	1,758	1.828	0.183	1.584	2.232
overhead efficiency, 1932	1,760	0.830	0.065	0.731	0.902
overhead efficiency (winsorized), 1932	1,760	0.832	0.052	0.731	0.902
Δ Democrat share of two-party vote, 1932–1936	1,763	-0.011	0.078	-0.129	0.119
% employed in agriculture, 1930	1,766	0.411	0.204	0.067	0.692
% employed in manufacturing, 1930	1,766	0.126	0.118	0.016	0.372
% white, 1930	1,766	0.973	0.066	0.888	1.000
% foreign-born, 1930	1,766	0.085	0.066	0.006	0.215
union potential, 1930	1,766	0.144	0.072	0.054	0.276
retail sales per capita, 1929	1,757	654.408	238.798	298.375	1,062.728
Δ retail sales per capita, 1933–1935	1,757	126.585	92.468	6.827	290.665
bank deposits per capita, 1932	1,743	0.142	0.160	0.013	0.344
Δ bank deposits per capita, 1932–1936	1,743	0.042	0.091	-0.036	0.130
IWW strikes per capita $\times$ 10,000	1,777	0.134	3.122	0.000	0.105
IWW locals per capita $\times$ 10,000	1,777	0.074	0.619	0.000	0.300
farms per capita, 1930	1,774	0.093	0.048	0.013	0.167
agricultural output per capita, 1930	1,767	155.323	132.464	11.463	387.508
unemployment rate, 1930	1,769	0.044	0.028	0.008	0.098
illiteracy rate, 1930	1,777	0.024	0.034	0.004	0.068
Δ Democrat share of two-party vote, 1928–1932	1,763	0.222	0.083	0.073	0.349
log agglomeration, 1930	1,758	0.601	0.106	0.460	0.803
log overhead efficiency, 1932	1,759	-0.190	0.084	-0.312	-0.103
log share of spending on highways, 1932	1,748	-2.055	0.679	-3.294	-1.147
log police and fire spending per capita, 1932	1,673	-7.151	1.192	-9.035	-5.175
log police and firemen per capita, 1932	1,659	-7.382	0.830	-8.842	-6.058
share of spending on highways, 1932	1,760	0.154	0.088	0.036	0.317
police and fire productivity, 1932	1,579	1.296	0.967	0.252	3.103

	Δ% Democrat 1932–1936									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
% urban	0.029**	0.034***	0.031***	0.034***	0.029**					
	(0.013)	(0.011)	(0.010)	(0.011)	(0.012)					
agglomeration						0.055**	0.063***	0.055***	0.062***	0.055***
						(0.020)	(0.019)	(0.019)	(0.018)	(0.017)
Income controls	x				x	x				x
Union controls		x			x		x			x
Additional controls			x		x			x		x
Δ% Democrat, 1928–1932				x	x				x	x
DV mean	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011
N	1729	1761	1761	1761	1729	1723	1754	1754	1754	1723
R^2	0.600	0.582	0.597	0.583	0.615	0.600	0.582	0.596	0.583	0.615

This table shows the results of regressions of the change in the Democrat share of the two-party vote against urbanization and agglomeration. All models include state fixed effects and controls for the shares in agriculture and manufacturing, the white and immigrant shares, and union potential. Income controls are retail sales per capita in 1929, the change in retail sales per capita between 1933 to 1935, bank deposits per capita in 1932, and the change in deposits per capita between 1932 and 1936. Union controls are the number of recorded IWW locals and strikes divided by population. Additional controls are farms and agricultural output per capita in 1930, and the unemployment and illiteracy rates in 1930. Models (4), (5), (9) and (10) control also for the change in the Democrats' change of the two-party vote between 1928 and 1932. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-7: Effects of urbanization on voting, controlling for income changes, unionization, development, and trends in voting

	Total		Loans		Grants		Relief Grants	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% urban	−0.110	−0.114	0.250	0.728***	−0.267***	−0.406***	0.457***	0.072
	(0.093)	(0.130)	(0.165)	(0.153)	(0.078)	(0.146)	(0.078)	(0.166)
Controls		x		x		x		x
DV mean	5.447	5.447	4.18	4.18	5.054	5.054	4.168	4.168
N	1758	1758	1756	1756	1758	1758	1758	1758
R^2	0.470	0.495	0.360	0.449	0.442	0.457	0.342	0.370

This table shows the results of county-level regressions of log New Deal spending divided by 1930 population against the urban population share. In (1) and (2) the dependent variable is log total New Deal spending per capita, in (3) and (4), New Deal loans, in (5) and (6), grants, and in (7) and (8), grants for relief. All models include state fixed effects, even-numbered models control for employment in agriculture and manufacturing, the white and foreign-born population shares, and union potential. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-8: Null relationship between urbanization and New Deal spending per capita

	Δ% Democrat 1932–1936				
	(1)	(2)	(3)	(4)	(5)
overhead efficiency	0.111**	0.103**	0.106***	0.108**	0.120***
	(0.041)	(0.039)	(0.038)	(0.040)	(0.039)
Income controls	x				x
Union controls		x			x
Additional controls			x		x
Δ% Democrat, 1928–1932				x	x
DV mean	-0.011	-0.011	-0.011	-0.011	-0.011
N	1727	1754	1754	1754	1727
R^2	0.603	0.585	0.600	0.587	0.618

This table shows the results of regressions of the change in the Democrat share of the two-party vote against the overhead efficiency. All models include state fixed effects and controls for the shares in agriculture and manufacturing, the white and immigrant shares, union potential, and the urban share. Income controls are retail sales per capita in 1929, the change in retail sales per capita between 1933 to 1935, bank deposits per capita in 1932, and the change in deposits per capita between 1932 and 1936. Union controls are the number of recorded IWW locals and strikes divided by population. Additional controls are farms and agricultural output per capita in 1930, and the unemployment and illiteracy rates in 1930. Models (4) and (5) control also for the change in the Democrats' change of the two-party vote between 1928 and 1932. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-9: Robustness of effects of overhead efficiency on voting, controlling for income changes, unionization, development, and trends in voting

	Δ% Democrat 1932–1936					
	(1)	(2)	(3)	(4)	(5)	(6)
agglomeration	0.146***	0.039**				
	(0.014)	(0.018)				
overhead efficiency			0.153***	0.091***	0.058**	0.080**
			(0.029)	(0.033)	(0.023)	(0.032)
% urban					0.110***	0.034***
					(0.008)	(0.010)
Controls		x		x		x
DV mean	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011
N	1754	1754	1756	1754	1754	1754
R^2	0.517	0.579	0.424	0.581	0.530	0.584

This table shows the results of regressions of the change in the Democrat share of the two-party vote against agglomeration and the overhead efficiency, not winsorized. All models include state fixed effects, even models include controls for the shares in agriculture and manufacturing, the white and immigrant shares, and union potential. Models (5) and (6) also control for the 1930 urban share. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-10: Robustness of effects of agglomeration and overhead efficiency on voting, not winsorizing

	Δ% Democrat 1932–1936			
	(1)	(2)	(3)	(4)
highway budget share	−0.224***	−0.115***		
	(0.047)	(0.030)		
police & fire productivity			0.022***	0.004**
			(0.002)	(0.002)
Controls		x		x
DV mean	-0.011	-0.011	-0.009	-0.009
N	1756	1754	1579	1579
R^2	0.446	0.585	0.485	0.585

This table shows the results of regressions of the change in the Democrat share of the two-party vote against alternative measures of productivity. The highway budget share is the share of government spending allocated to highway maintenance. Police and fire productivity is the exponential of the residual from a regression of log police and fire employment per capita against log spending per capita on police and fire protection. All models include state fixed effects, even models include controls for the shares in agriculture and manufacturing, the white and immigrant shares, and union potential. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-11: Effects of alternative measures of productivity on realignment towards Democrats

	log spending					
	(1)	(2)	(3)	(4)	(5)	(6)
overhead efficiency	6.906***	4.606***	3.673***	−0.140	−0.053	−0.051
	(0.770)	(0.669)	(0.535)	(0.285)	(0.316)	(0.314)
% urban			2.061***			0.066
			(0.217)			(0.055)
log population				0.970***	0.918***	0.911***
				(0.021)	(0.020)	(0.020)
Controls		x	x		x	x
DV mean	6.224	6.224	6.224	6.224	6.224	6.224
N	1760	1758	1758	1760	1758	1758
R^2	0.473	0.713	0.770	0.963	0.968	0.968

This table shows the results of county-level regressions of log spending against overhead efficiency. The dependent variable is log of government spending exclusive of highway maintenance in 1932. All models include state fixed effects, models (2), (3), (5), and (6) control for the shares in agriculture and manufacturing, the immigrant and white population shares, and union potential. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-12: Relationship between overhead efficiency and 1932 spending, for non-Southern counties

Table A-13: Effect of urban status on Gallup survey responses

	Base		Controls			
	Coef	S.E.	Coef	S.E.	Mean	N
Government Spending and Taxation						
Support increased government spending	0.123***	(0.017)	0.0755***	(0.0208)	0.388	3,665
Government spending too little on relief	0.116***	(0.0259)	0.0771***	(0.0249)	0.336	2,276
Preferred income tax rate (%)	0.0435*	(0.0232)	0.0842***	(0.0308)	0.237	1,884
Support for New Deal Democrats						
Prefer New Dealer to conservative Democrat	0.102***	(0.0221)	0.0745***	(0.0264)	0.578	3,173
Want more liberal policies from Roosevelt	0.113***	(0.0141)	0.0973***	(0.0146)	0.335	6,106
Prefer a Liberal to a Conservative party	0.0886***	(0.0155)	0.0565***	(0.0175)	0.487	7,712
Regard self as liberal	0.0187	(0.0249)	−0.000152	(0.0284)	0.515	2,828
Support for Specific Programs						
Approve of TVA	0.164***	(0.028)	0.137***	(0.0319)	0.568	1,292
Support increased spending on unemployment relief	0.178***	(0.0207)	0.128***	(0.0214)	0.385	2,412
Support increased spending on farm benefits	−0.00732	(0.0211)	0.00472	(0.023)	0.514	2,294
Support revival of AAA	−0.0733***	(0.0257)	−0.0575*	(0.032)	0.384	3,605
Support old age pensions	0.0531*	(0.0322)	0.0316	(0.0388)	0.705	3,484
Approve of Social Security	0.0882***	(0.0166)	0.0478**	(0.0195)	0.748	2,590
Approve of tax for Social Security	0.0909***	(0.024)	0.0508*	(0.0263)	0.733	2,348

This table shows the results of regressions of agreement with Gallup poll questions on urban status. Dependent variables are coded so that 1 indicates agreement with the statement and 0 indicates disagreement, with the exception of the tax question, which is in percentage points. The Base specification only includes survey-wave fixed effects. The Controls specification includes controls for region, race, occupation, gender, and survey-wave. All models are restricted to non-southern respondents and weighted using the population weights developed by Berinsky and Schickler. Robust standard errors in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

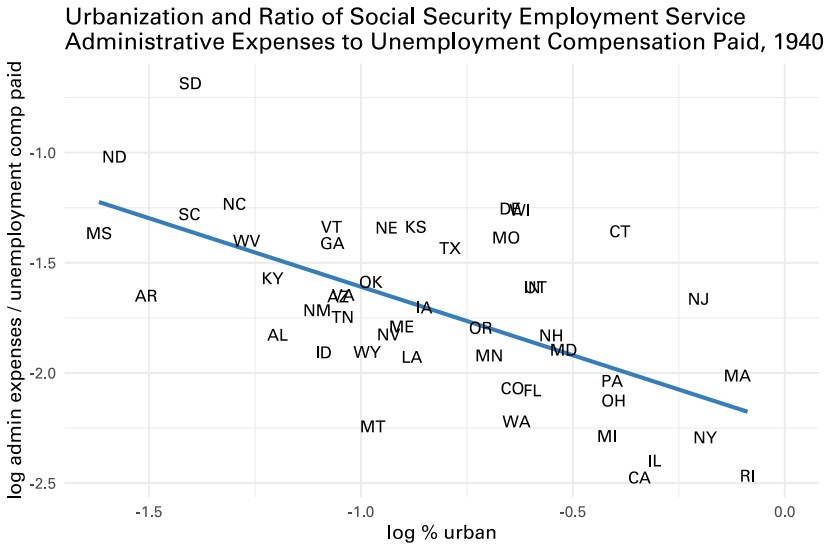

Figure A-3: Lower Social Security administrative costs relative to disbursements in more urban states

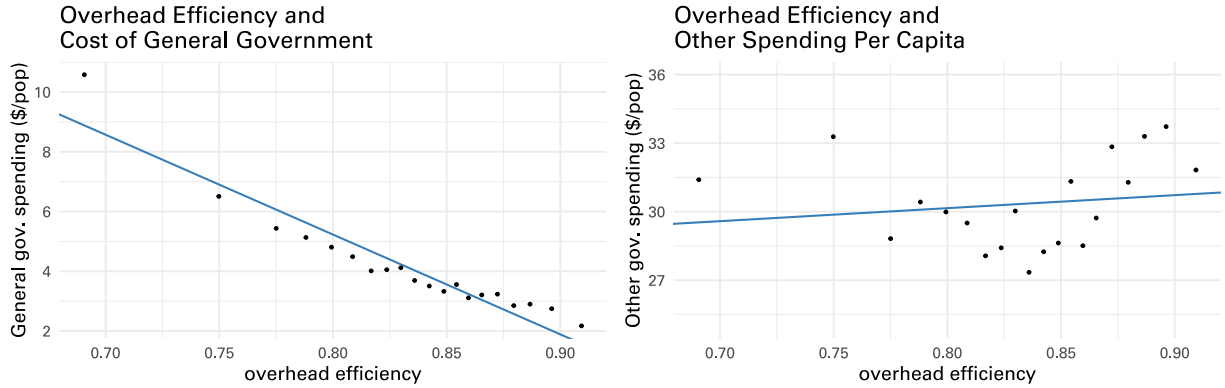

Figure A-4: Overhead efficiency correlates with lower government overhead spending per capita, not with higher spending in other policy areas

This figure shows binned scatterplots of the relationship between overhead efficiency at the county level and log general government overhead spending per capita (left) and log government spending per capita over other categories (right).

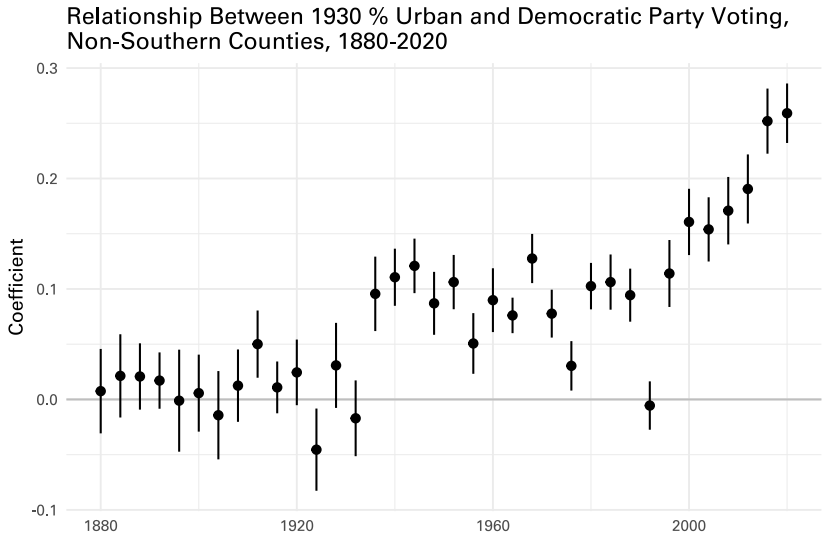

Figure A-5: Regression evidence for the emergence of the urban-rural divide in 1936, with county and state-by-year fixed effects

This figure shows coefficients and 95% confidence intervals from a regression of the Democratic share of the two-party vote against 1930 % urban, for non-Southern counties, controlling for county and state-by-year fixed effects

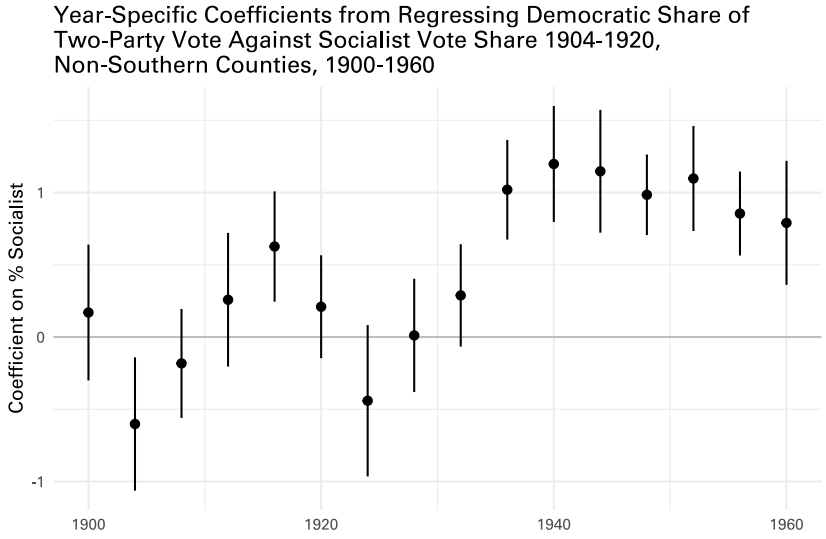

Figure A-6: The relationship between historical Socialist Party support and voting for the Democratic Party began in 1936

This figure shows coefficients and 95% confidence intervals from a regression of the Democratic share of the two-party vote against the average vote share for the Socialist Party over the period 1904–1920, for non-Southern counties, controlling for county and state-by-year fixed effects

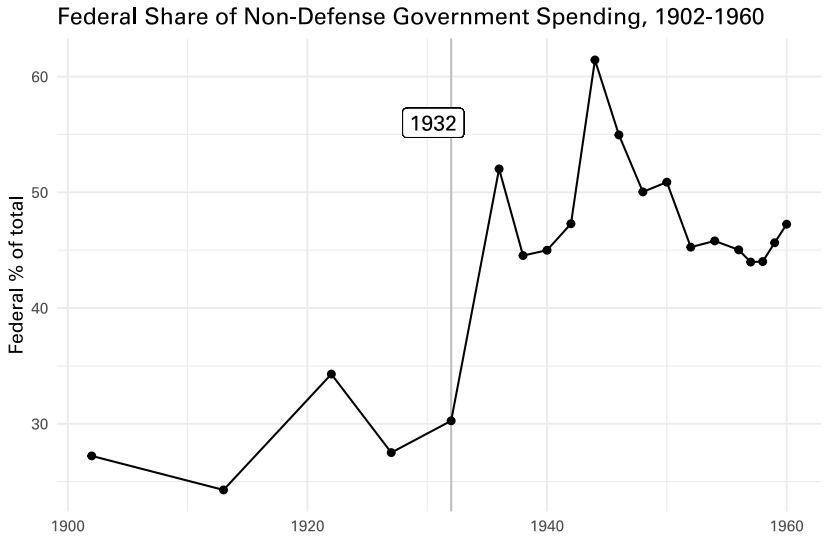

Figure A-7: The federal share of US government spending radically increased after 1932

This figure shows the share of government spending net of national defense and international relations attributed to the Federal Government, from US Department of Commerce (1969). Transfers between levels of government are subtracted from the recipient level.

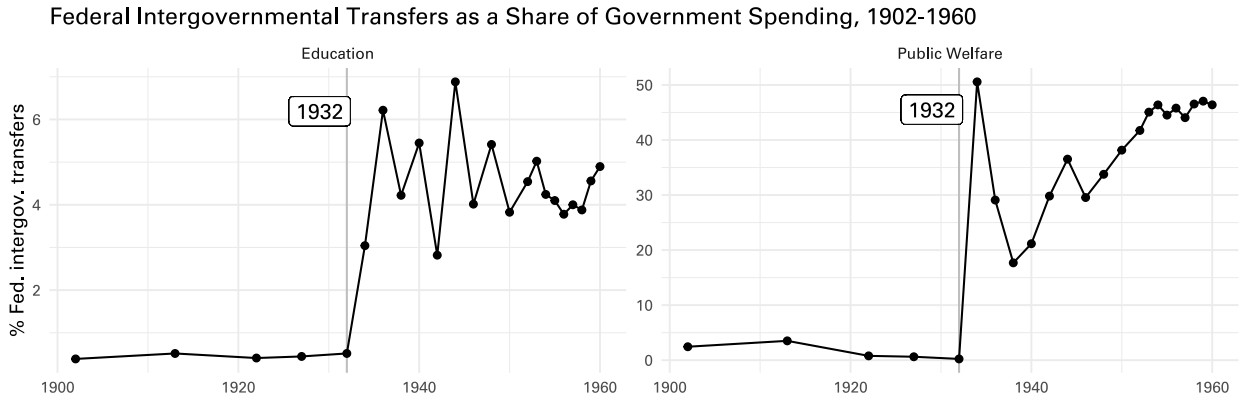

Figure A-8: Federal intergovernmental transfers for education and public welfare increased after 1932

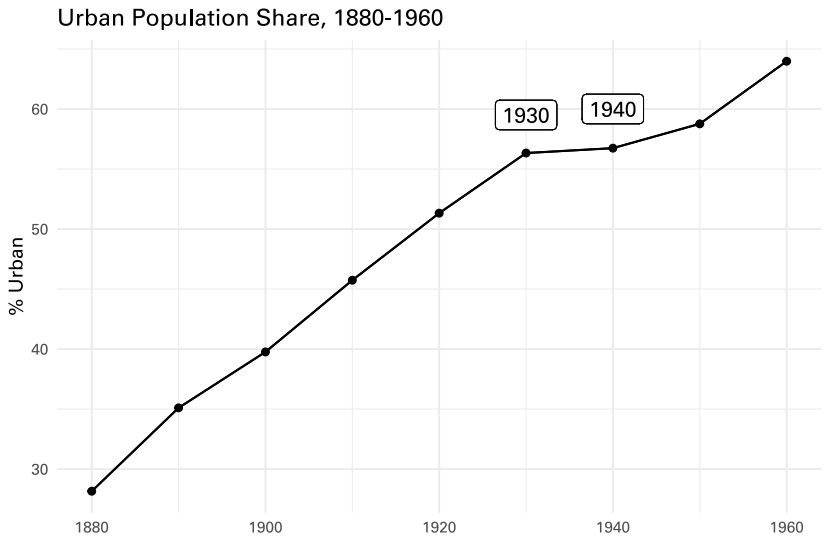

Figure A-9: The pace of urbanization was slower in the 1930s than preceding decades

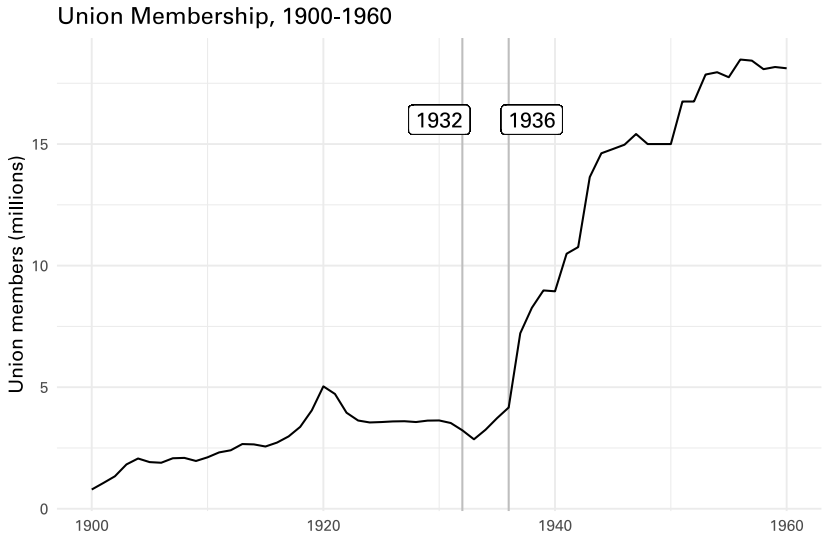

Figure A-10: The large increase in union membership took place after 1936

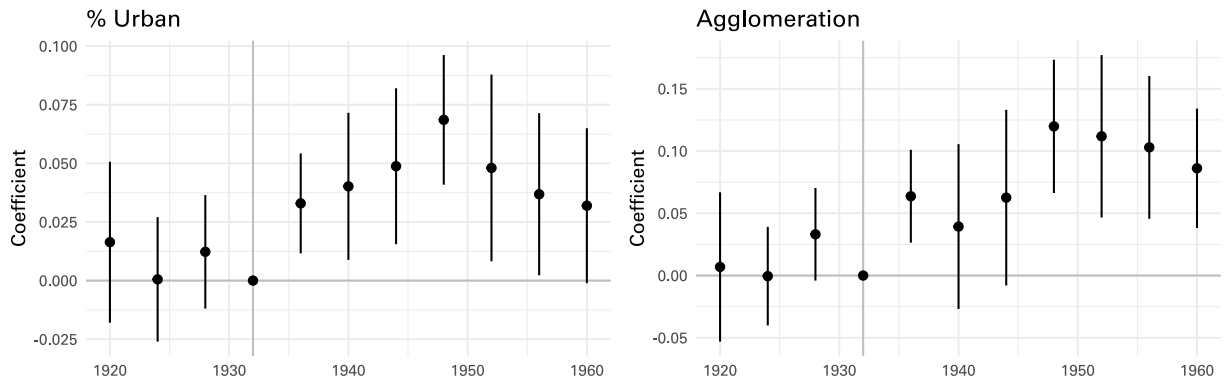

Figure A-11: Dynamic effects of urbanization on voting, with controls

These figures show coefficients and 95% confidence intervals from regressions of the Democratic share of the two-party vote against 1930 % urban and agglomeration interacted with year indicators, with county and state-by-year fixed effects and controls for employment in agriculture and manufacturing, the white and foreign-born shares of the population, and union potential, all interacted with year indicators.

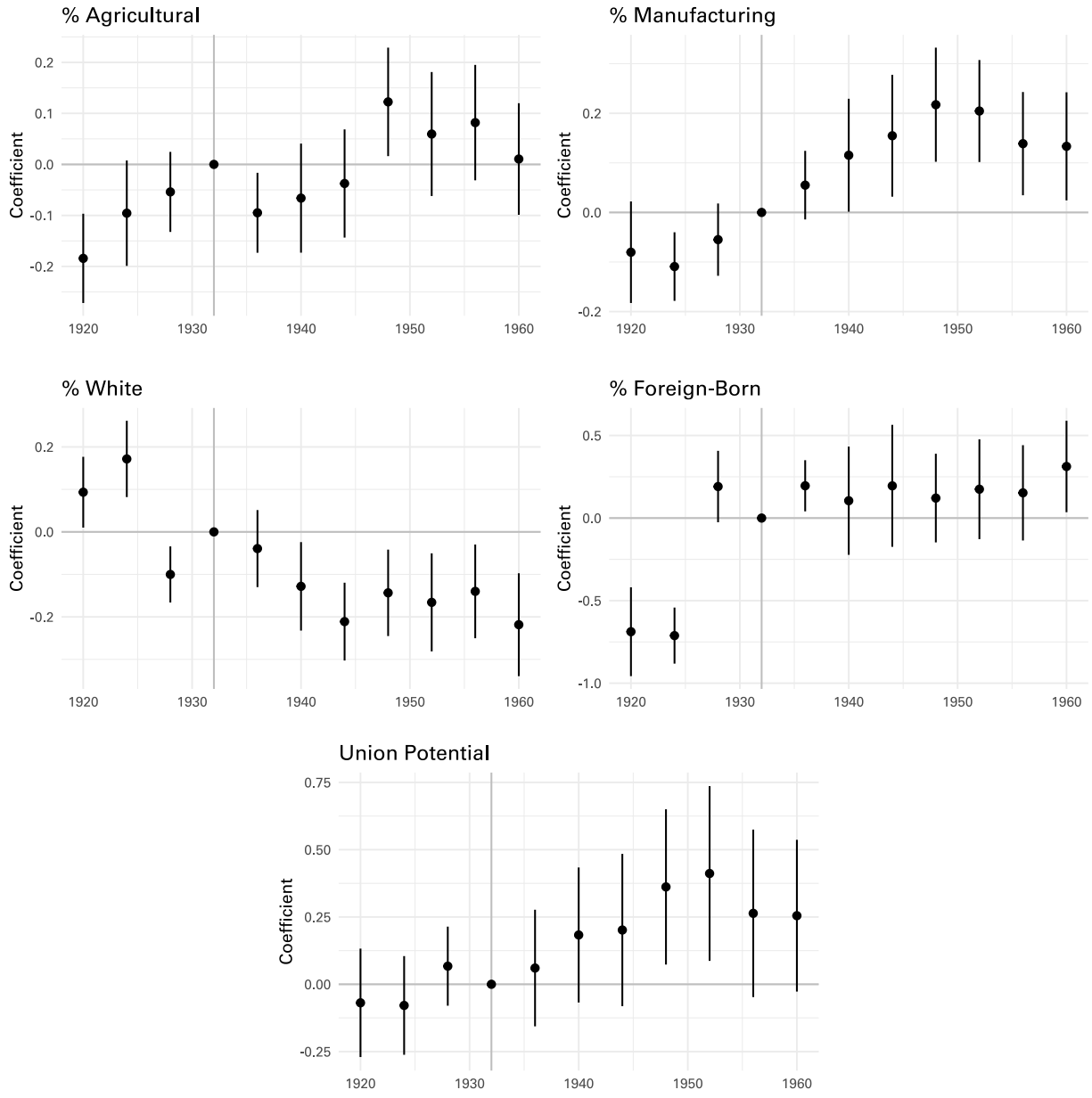

Figure A-12: Effects of control variables on the realignment, controlling for all others

This panel shows results of regressing the county-level Democrat share of the two party vote against different economic variables measured in the 1930 census interacted with year indicators, with county and state-by-year fixed effects. Each plot shows the coefficients on a given variable, from the same model used in the left panel of Figure A-11 controlling for the other four variables and 1930 % urban interacted with year indicators, with 95% confidence intervals calculated using standard errors clustered by state.

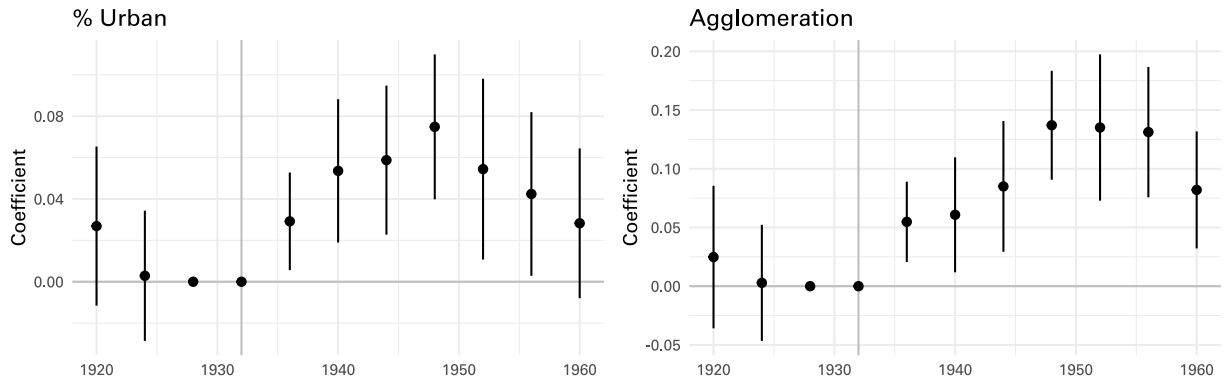

Figure A-13: Dynamic effects of agglomeration on voting, with full set of controls from Table A-7

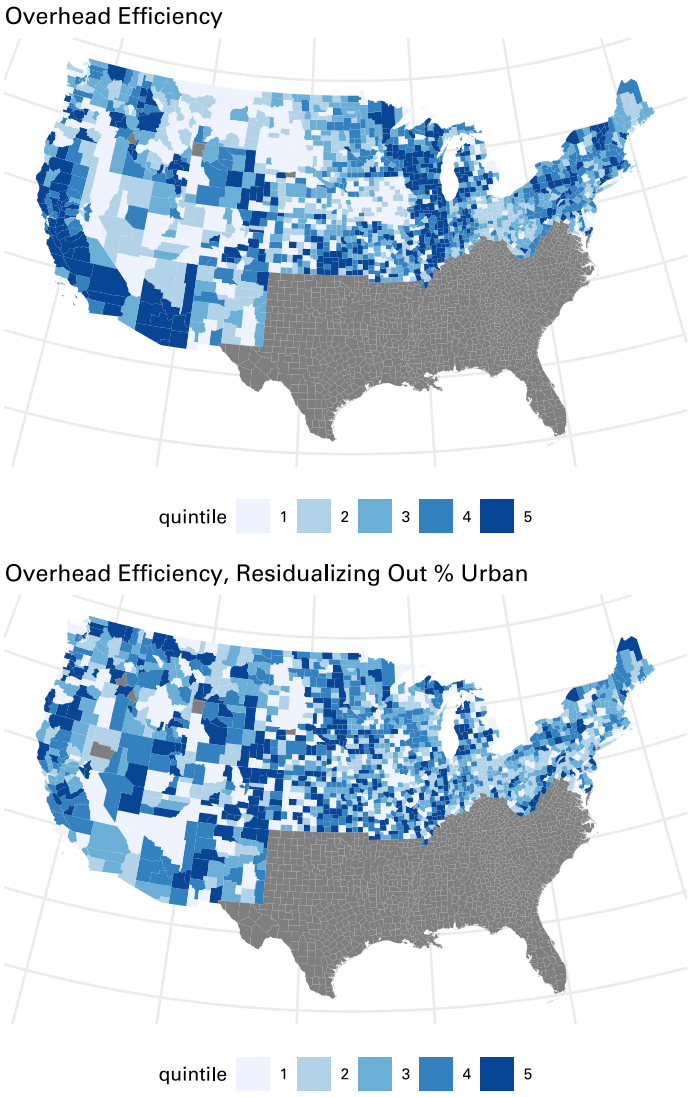

Figure A-14: Spatial distribution of overhead efficiency

The top panel maps overhead efficiency, the bottom maps the residuals from regressing overhead efficiency against 1930 % urban and state fixed effects

Table A-14: Gallup Poll Questions and Sources

Question	Full Question or Questions	Source
Support increased government spending	Do you think government spending should be increased to help get business out of its present slump? (1 = YES!, 0 = NO!)	Gallup Poll # 1938-0120, April 21–26, 1938; # 1938-0123, May 22–27, 1938
Government spending too little on relief	In your opinion, is the government spending too little, too much, or the right amount for relief and recovery? (1 = Too Little, 0 = Too Much)	Gallup Poll # 1937-0077, April 7–12 1937
Preferred income tax rate (%)	How much do you think a married man earning $3,000/$5,000/$10,000 a year should pay in the form of income taxes (answer as % of income)	Gallup Poll # 1937-0077, April 7–12, 1937
Prefer New Dealer to conservative Democrat	If Roosevelt is not a candidate for reelection in 1940 would you prefer a conservative type of candidate, or a New Dealer (88, 109); Asked of Democrats, If Roosevelt is not a candidate in 1940, would you prefer a conservative or a New Dealer for President? (126) (1 = New Dealer, 0 = Conservative)	Gallup Poll # 1937-0088, June 23–28, 1937; # 1938-0109, January 20–25, 1938; # 1938-0126, June 23–28, 1938
Want more liberal policies from Roosevelt	Should President Roosevelt's second administration be more liberal, more conservative, or about the same as his first? (56) (1 = More Liberal, 0 = More Conservative); Do you think the policies of the Roosevelt administration are too liberal, too conservative, or about right? (103, 104) (1 = Too Conservative, 0 = Too Liberal), During the next two years would you like to see the Roosevelt Administration be more liberal or more conservative (125) (1 = More Liberal, 0 = More conservative)	Gallup Poll # 1936-0056, November 6–11, 1936; # 1937-0103, November 14–19, 1937; # 1937-0104, November 21–26, 1937; # 1938-0125, June 11–16, 1938
Prefer a Liberal to a Conservative party	If there were only two political parties in this country—one for conservatives and one for liberals—which would you join? (69), If these two new parties were formed [Liberals and Conservatives], which one do you think you would join? (94); Which party [Liberals and Conservatives] do you think you would like to join? (118); If there were only two political parties in this country—one for conservatives and the other for liberals—which do you think you would join? (127); If the Democratic and Republican parties went out of existence would you join the Conservative or the Liberal party? (132) (1 = Liberal, 0 = Conservative)	Gallup Poll # 1937-0069, February 17–22, 1937; # 1937-0094, August 4–9, 1937; # 1938-0118, April 8–13, 1938; # 1938-0127, July 4–9, 1938; # 1938-0132, September 15–20, 1938
Regard self as liberal	In politics, do you regard yourself as a radical, a liberal, or a conservative? (76); In politics, do you regard yourself as a liberal or a conservative? (109) (1 = Liberal, 0 = Conservative)	Gallup Poll # 1937-0076, April 1–6, 1937; # 1938-0109, January 20–25, 1938

Approve of TVA	Do you approve of the TVA? (1 = YES!, 0 = NO!)	Gallup Poll # 1938-0120, April 21–26, 1938
Support increased spending on unemployment relief	Do you think government expenditures should be increased or decreased on the following: Unemployment relief? (1 = Increased, 0 = Decreased)	Gallup Poll # 1937-0101, October 20–25, 1937
Support increased spending on farm benefits	Do you think government expenditures should be increased or decreased on the following: Farm benefits? (1 = Increased, 0 = Decreased)	Gallup Poll # 1937-0101, October 20–25, 1937
Support revival of AAA	Would you like to see the AAA (crop control act) revived? (1 = Yes, 0 = No)	Gallup Poll # 1937-0068, February 10–15 1937; # 1937-0069, February 17–22, 1937
Support old age pensions	Do you favor the compulsory old age insurance plan, starting in January, which requires employers and workers to contribute equally towards workers' pensions? (56); Do you favor the present compulsory old age insurance plan which requires employers and workers to contribute equally toward workers' pensions? (65) (1 = Yes, 0 = No)	Gallup Poll # 1936-0056, November 6–11, 1936; # 1937-0065, January 20–25, 1937
Approve of Social Security	Do you approve of the present Social Security laws which provide old age pensions and unemployment insurance? (1 = YES!, 0 = NO!)	Gallup Poll # 1938-0127, July 4–9, 1938
Approve of tax for Social Security	Do you approve of this tax? [to fund social security] (1 = Yes, 0 = No)	Gallup Poll # 1938-0107, December 30, 1937–January 4, 1938

	Δ% Democrat 1932–1936					
	(1)	(2)	(3)	(4)	(5)	(6)
% urban	0.087***	0.035***	0.103***	0.037***	0.111***	0.034***
	(0.009)	(0.009)	(0.006)	(0.010)	(0.008)	(0.010)
1940 migrant % urban	0.094***	0.002				
	(0.031)	(0.039)				
1940 destination average county % urban			0.054	−0.042		
			(0.036)	(0.035)		
1940 destination average overhead efficiency					0.204	0.196
					(0.223)	(0.233)
Controls		x		x		x
DV mean	-0.011	-0.011	-0.011	-0.011	-0.011	-0.011
N	1761	1761	1761	1761	1761	1761
R^2	0.536	0.582	0.531	0.582	0.530	0.582

This table shows the results of county-level regressions of the voting against urbanization and measures of migration to urban locations. The dependent variable is the change in two-party vote share for the Democrats, 1932–1936. "1940 destination % urban" is the share of those listed as present in the county in 1935 living in an urban place in 1940, "1940 destination average county % urban" is the weighted average urban population share of counties in which people present in the county in 1935 were living in 1940, "1940 destination average overhead efficiency" is the weighted average overhead efficiency of these counties. All models include state fixed effects, even models add controls for the shares in agriculture and manufacturing, the immigrant and white population shares, and union potential. Standard errors clustered by state in parentheses. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A-15: Future migration to cities is not associated with voting Democrat

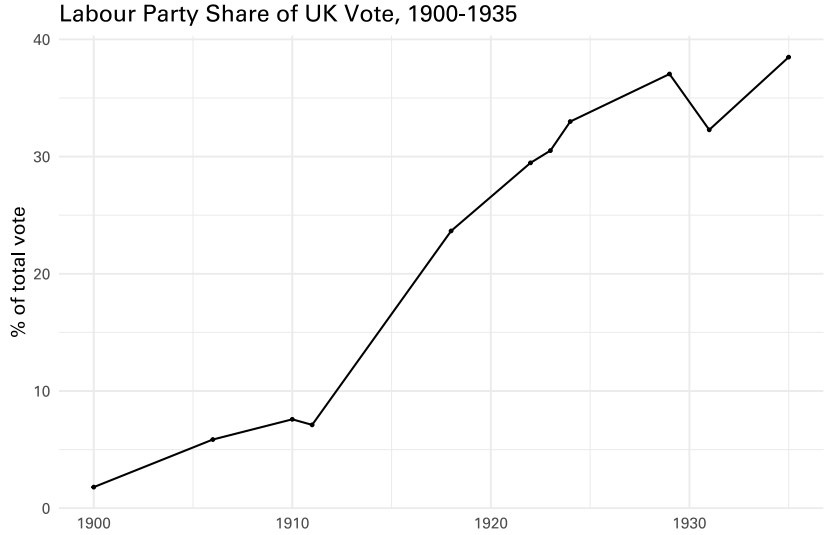

Figure A-15: The rise of Labour in British politics