



Is strategic interaction among governments just a modern phenomenon? Evidence on welfare competition under Britain's 19th-century Poor Law

Jan K. Brueckner¹

Accepted: 3 December 2021 / Published online: 25 January 2022
© The Author(s) 2022

Abstract

Drawing on data from mid-19th century Britain, this paper studies strategic interaction among local governments in the choice of welfare benefits under the Poor Law, the local welfare system of the time. The paper exploits a national reform that reduced the length of residency required for welfare eligibility, which should have increased the incentive for welfare migration and thus led to both stronger strategic interaction and lower levels of equilibrium spending. The results show evidence of a positive but small degree of baseline interaction, suggesting that modern models of welfare competition may apply even in settings with relatively high migration costs. While the change in post-reform equilibrium spending is negative as predicted, the results show no evidence of stronger interaction after the reform.

Keywords Poor Law · Welfare competition · England

JEL Classification H72 · N93

1 Introduction

A large literature has emerged studying strategic interaction among governments. The literature includes many theoretical papers, most of which focus on tax competition as a source of strategic interaction (see Wilson, 1999 for a survey). Spurred by this theoretical research, empirical work on strategic interaction has flourished. The goal of most empirical papers is to estimate policy reaction functions, which relate the level of a jurisdiction's policy variable (a tax rate, for example) to the policy choices of its neighbors (see Brueckner, 2003; Revelli, 2005 for surveys). A

✉ Jan K. Brueckner
jkbrueck@uci.edu

¹ Department of Economics, University of California, Irvine, Irvine, CA 92697, USA

nonzero slope coefficient in the estimated reaction function is evidence of strategic interaction.

While most empirical studies focus on tax rates, a smaller group investigates strategic interaction in spending levels. Within this latter set, a handful of papers looks for strategic interaction in the choice of welfare-benefit levels for poor households. Such interaction, known as “welfare competition,” is the mirror image of tax competition, with high benefit levels attracting poor migrants via “welfare migration,” but with high taxes repelling business investment and mobile taxpayers. In both cases, however, the result is an inefficiently low level of the policy variable, welfare benefits or tax rates. The three existing papers focusing on welfare competition are Figlio et al. (1999), Saavedra (2000), and Dahlberg and Edmark (2008), each of which finds evidence of strategic interaction in the choice of welfare benefits. The present paper adds to this small literature.

Given the exhaustive overall body of empirical work on strategic interaction, any new paper faces a high hurdle in justifying its existence. Such a paper, whether its focus is on tax rates, welfare benefits, or some other policy variable, should offer a distinct methodological advancement, or other unique features not found in the previous literature. The paper clears this hurdle in part by offering the first study of strategic interaction using historical data. Relying on data from mid-19th century Britain, the paper focuses on the choice of welfare benefits by British districts under the New Poor Law, the welfare system of the time. Under this system, applicants were eligible for assistance partly on the basis of the length of their residency, a requirement that was made less stringent during the sample period. Beyond its historical focus, the paper is further distinguished from previous work by its investigation of the effects of this length-of-residency reform, motivated by theoretical predictions.

Specifically, the analysis exploits the Irremovable Poor Act of 1861, which was passed in the midst of the sample period and reduced the residency requirement for welfare eligibility from five to three years. This change likely made welfare migration more attractive, and thus had the potential to alter the nature of strategic interaction among districts. The paper’s theoretical model predicts that under a plausible assumption, the looser residency requirement should have led to stronger interaction among districts (a steeper reaction function), while also reducing the equilibrium level of welfare benefits. To the best of our knowledge, the current tests of these hypotheses have no similar precedent in the strategic-interaction literature. The study thus makes an additional contribution by shedding light on the sensitivity of local welfare policy to changes in migration incentives.¹

It is useful to note that in the USA, various short-lived welfare residency requirements imposed at the state level were all declared unconstitutional decades ago, precluding a similar study of their effect on state benefit levels. But the strategic interaction uncovered by Figlio et al. (1999) and Saavedra (2000) in their US studies

¹ Agrawal et al. (2021) offer another recent innovation in the analysis of strategic interaction by investigating the effect of cooperative agreements among French municipalities on the intensity of tax competition, finding that cities with such agreements compete less intensely than those without them. See also Eugster and Parchët (2019).

appears to confirm that the migration incentives these residency requirements were meant to undercut were indeed at work in the USA, driving state benefit decisions.² What new would be learned by showing that the same forces were operating in 19th-century Britain? On the one hand, such results would bolster the external validity of modern models of local government behavior, showing that they are broadly applicable even to an industrializing society. However, given that greater migration frictions presumably existed at that time, the emergence of strategic interaction might even be viewed as surprising, showing that even relatively limited mobility was enough to generate welfare competition among British districts.

The pre- and post-reform structure of the sample period leads to an empirical model where the reaction functions are allowed to differ across the two subperiods. It is then possible to gauge the baseline extent of interaction among districts, along with short-run change in interaction as residency requirements are relaxed, a novel design. The empirical setup, however, is more conventional in another respect. In particular, to address the endogeneity of welfare benefits of neighboring districts, which appear on the right-hand side of the reaction function, the regressions rely on instrumental variables that are traditional in the literature: the demographic and other characteristics of neighboring jurisdictions, which help to determine the neighbors' benefit levels.³

In addition to contributing to research on strategic interaction among governments, the paper also advances the literature analyzing England's economic and political institutions in the industrial revolution era. Following a heyday between 1970 and 1990 [see references in Boyer (2002), and studies based on Southall et al., (1998), economic research on the evolution of the 19th-century British welfare state has seen a recent resurgence, with studies including Chapman (2018, 2020a) examining the causes and consequences of public spending, and other work including Boyer (2019) and Horrell et al. (2020) documenting its impact on living standards and economic development. As in these more recent studies, the present paper offers insights into debates among contemporary officials and observers, who argued about whether poor relief influenced migration and whether relief generosity responded to migration (Boyer, 1993).⁴ While earlier studies were often based on limited data,

² A sizable past empirical literature looked for evidence of welfare migration in the US. See Blank (1988) for one of the best such studies and Brueckner (2000) for a survey of the remaining literature.

³ While some recent methodological critiques (Gibbons and Overman (2012), for example) have argued that neighbor characteristics may not necessarily be excludable as covariates in the reaction function, making them illegitimate as instruments for the neighbor policy variable, diagnostic evidence is provided in support of the validity of the traditional instruments, lending credence to the current results. Note as well that the same identification issues described in Gibbons and Overman (2012) also arise in the models of peer effects, which are formally similar to models of government interaction. See for example, Patacchini and Zenou (2012). As an alternative, recent papers, including Dahlberg and Edmark (2008), Baskaran (2014), Lyytikäinen (2012), Agrawal (2015), Parchët (2019), exploit institutional features of their data environments to generate different instruments that are less subject to this excludability criticism. See below for further discussion.

⁴ The paper also adds to a small literature examining the role of relief spending in migration patterns in other historical settings. See, for example, Boustan et al. (2010).

the use of newly collected, high-resolution data on poor relief covering the entire country allows this paper to provide better evidence bearing on these debates.

The plan of the paper is as follows. Section 2 describes the historical context of the study, and Sect. 3 presents a stylized theoretical model that generates the two testable hypotheses described above. Section 4 derives the estimating equation and discusses identification. Section 5 discusses the data, and Sect. 6 presents the empirical results. Section 7 discusses interaction in the two components of welfare spending (indoor and outdoor relief), and Sect. 8 offers conclusions.

2 Historical context

The welfare system known as the Old Poor Law was established in England and Wales by the Poor Relief Act of 1601, which created a right to relief at the national level that was implemented locally, with relief funds raised and distributed at the parish level. Spatial variation in relief levels under the Old Poor Law appeared to create a migration incentive for the poor, even though “the vast majority” of moves of any type occurred over short distances (Lloyd, 2007). In particular, as parishes adjusted to the new Act, some “were more sympathetic toward their poor” than others, with paupers moving in “from less generous parishes,” as noted by Bloy.⁵ The Settlement Act of 1662 was designed to limit such welfare migration, requiring an individual to be “settled” in order to receive relief from a parish, which meant being born in the parish or continuously employed there for a year, a condition hard to satisfy because of intentionally short labor contracts (Bloy).

By the early 1800s, around 10% of the English population was receiving relief in some amount under the Old Poor Law (Clark and Page, 2019), leading observers to believe that the system was being abused (Fraser, 2009). In response, the Poor Law Amendment Act of 1834 created the New Poor Law, which continued to mandate local redistribution, though with restrictions to limit abuse. It was the main source of poor relief in mid-to-late 19th-century England and Wales.

Under this system, a national Poor Law Commission (later, the Poor Law Board) regulated the broad features of welfare policy, while collections of parishes called Poor Law Unions were responsible for administering and financing the relief system (Boyer, 2002). Poor Law Unions had considerable latitude in setting the generosity of relief payments and the tax rates on land and property that supported them. Through the 1834 Act, the local landowners whose tax payments financed the system gained greater power than before in setting the levels of relief, and one result was a 30% decline in spending between 1834 and 1861 (Chapman, 2020b).⁶ A

⁵ This quote is from Marjorie Bloy’s webpage on the Old Poor Law under “The Web of English History” (<http://www.historyhome.co.uk/peel/poorlaw/plaa.htm>).

⁶ It is possible that this decline served a national goal by inducing migration out of poorer rural areas into the industrializing cities. Also note that, like the present paper, Chapman, (2020b) uses Poor Law data to ask a focused question. He explores the effect of a 1894 reform that lessened landowners’ influence over relief levels, showing that spending rose subsequently in Poor Law Unions with the greatest degrees of inequality.

convenient empirical feature of Poor Law Unions is their close match to the boundaries of registration districts, which were used for census purposes and constitute the unit of observation in the empirical analysis.

Lindert (1998) estimates that over the 1850s and 1860s, total relief expenditure under the New Poor Law made up roughly 0.86–1.07% of GDP, having fallen from 2.7% in 1820–21, a level not reached again until late in the century (Feldman, 2003). The two main categories of public assistance under the system were outdoor relief (direct cash payments to the poor), and indoor relief. Under this second category, paupers were provided with shelter and subsistence diets inside “workhouses,” where labor efforts were required in return for these benefits. Labor often consisted of “tasks such as breaking stones, crushing bones to produce fertilizer, or picking oakum,” which typically generated revenue insufficient to cover the workhouse’s operating costs.⁷ Workhouses were designed to be restrictive and unpleasant so as to deter voluntary unemployment, and the shelter cost they entailed made indoor relief the more expensive of the two types.⁸ Outdoor relief, in addition to supporting the unemployed, was often used to supplement labor income for non-destitute individuals earning low wages from employment, and it also contained payments for children.

While the New Poor Law, reflecting concerns about abuse, mandated the elimination of outdoor relief for the able-bodied poor (who would be supported in workhouses), the use of outdoor relief persisted. Out of 1.33 million paupers in 1846, 375,000 able-bodied poor received outdoor relief, while only 82,000 were in workhouses (Fraser, 2009). The split of spending between indoor and outdoor relief was another element over which Poor Law Unions exercised discretion.

Eligibility for both types of assistance was determined by a combination of the applicant’s physical ability to work (gauged via “labor tests”) and their “rights of settlement,” as noted above (Boyer, 1993; Rose, 1976). While “unsettled” individuals, who were typically born elsewhere, might have benefited from relief, they were subject to “removal,” or deportation to the true place of settlement, if they sought relief or even seemed likely to seek it (Feldman, 2003). A contemporary source quoted by Feldman stated that the threat of removal was “hung up *in terrorem* over the heads of the poor” to deter them from applying for relief (p. 90).

Even though poor residents were disadvantaged by the decline in relief levels after 1834, the threat of removal was lessened by the passage of the Poor Removal Act of 1846, which allowed residents to become eligible for relief after living continuously in a district for 5 years. The removal threat was reduced further by a second national reform, the Irremovable Poor Act of 1861, which reduced the residency requirement from 5 to 3 years (Feldman, 2003; Rose, 1976). Feldman states that

⁷ This quote and the operating-cost claim are from the Wikipedia entry “Workhouse” (https://en.wikipedia.org/wiki/Workhouse#cite_note-FOOTNOTEFraser200967-30). Picking oakum is the task of preparing the tarred fibre used in sealing gaps, as in caulking of ship planks. The Wikipedia article also argues that some private firms with outputs matching those of workhouses viewed the institutions as a competitive threat.

⁸ Poor houses were thus a form of “workfare,” which is shown by Besley and Coate (1992) to be superior under some conditions to unrestricted cash welfare grants.

“these new acts forced urban authorities and urban ratepayers to take responsibility for the welfare of their migrant poor in ways hitherto they had been able to evade” (p. 93). A district’s discretion also extended to the enforcement of these residency restrictions, with lax enforcement possibly lessening the removal threat.

As explained in the introduction, one goal of this paper is to study the response of local Poor Law authorities to the weakening of the residency requirement following the Irremovable Poor Act of 1861. With a weaker requirement, welfare migration may have been a bigger threat following the reform, potentially changing the strength of strategic interaction among districts as they set benefit levels.⁹ The absence of other significant changes in the British economic environment in the time window around the reform allows reliable measurement of this response.

3 Theoretical model

This section develops a stylized theoretical model of welfare competition to motivate the subsequent empirical work. The framework is adapted from Brueckner (2000), who in turn builds on Wildasin (1991).¹⁰ In the model, the district’s concave production function is $f(n)$, where n is the labor input. Workers are poor, receiving a welfare benefit (if eligible), in addition to a low wage w . Assuming f is quadratic, the wage (equal to the marginal product f') is linear in n , given by $w = \alpha - \beta n$, where $\alpha, \beta > 0$.

For simplicity, the analysis focuses on the case with just two districts, 1 and 2. Workers can migrate between districts to secure a higher effective income, equal to the wage plus the effective welfare payment. This effective payment is equal to the statutory welfare benefit, E , times an adjustment factor $a < 1$ that captures the welfare residency requirement. With a worker needing to live in the district for several years before getting benefits, the effective benefit for a new migrant is lower than E , being equal to aE to account for the initial years when it is not paid. Thus, a looser residency requirement, with fewer years of residence needed following the reform, corresponds to a larger value of a .

The welfare-migration equilibrium, which equates effective incomes of the poor between the two districts, then satisfies $w_1 + aE_1 = w_2 + aE_2$, or

⁹ Note that the 1861 act was followed by a second relaxation of residency requirements in 1865: among the provisions of the Union Chargeability Act of 1865 was a reduction in the residency period from 3 to 1 years. To allow a cleaner interpretation, and because the 1865 act also included substantial changes in the way Poor Law Unions were funded and in the balance of local power, the current analysis ends before this policy came into effect. Nevertheless, it is feasible to include both reforms by defining the pre-reform period as 1857–1860 and the post reform period as 1865–1870. While strategic interaction emerges under this setup, the results show an increase, rather than the predicted decrease, in welfare benefits following both reforms. Since the reverse outcome (a decrease in welfare benefits) occurs when attention is restricted to the first reform (as seen below), and since the two-reform approach may be contaminated by other factors, the latter approach is dispreferred.

¹⁰ See Wheaton (2000) for a related model.

$$\alpha - \beta n_1 + aE_1 = \alpha - \beta(N - n_1) + aE_2, \tag{1}$$

where N is the total worker population of the two districts (so that $n_2 = N - n_1$). Solving (1) for n_1 yields

$$n_1 = \frac{N}{2} + \frac{a(E_1 - E_2)}{2\beta}. \tag{2}$$

If $E_1 = E_2$, then workers divide equally between the districts, while if $E_1 > E_2$, district 1 has more workers than 2. Its wage w_1 is then lower than w_2 , which serves to equalize overall incomes. Note that a larger a accentuates the interdistrict population difference caused by a difference in the E 's.

Each district has M nonpoor property owners who finance the welfare benefits, and those in district 1 pay a per capita tax equal to an_1E_1/M . Letting y denote the common income of all property owners, their consumption level is then equal to

$$x_1 = y - \frac{an_1E_1}{M}. \tag{3}$$

The property owners value x consumption while also caring about the income of the poor, either because of altruism or concerns about poverty and social unrest. Suppose that, in evaluating the poor's income, the property owners focus on the statutory welfare benefit E_1 , not the effective benefit aE_1 , which is determined by the residency rules imposed nationally. Thus, property owners in district 1 care about the statutory income $I_1 = w_1 + E_1$ of the poor. Suppose that the owners have quasi-linear preferences, with I_1 entering in quadratic fashion. The utility of district 1's property owners is then written as

$$u(x_1, I_1) = x_1 + \delta I_1 - \theta I_1^2, \tag{4}$$

where $\delta, \theta > 0$ and I_1 again equals $w_1 + E_1$.

Substituting for x_1 in (4) using (3), and substituting for w_1 (part of I_1) using $w_1 = \alpha - \beta n_1$, utility is then a function of n_1 and E_1 . Substituting further for n_1 using the solution from (2), utility in (4) reduces to an expression that depends on E_1 and E_2 .¹¹ District 1's property owners maximize this expression by choice of E_1 viewing E_2 as parametric. The solution for E_1 yields district 1's reaction function (or best

¹¹ The utility expression and the expression for poor income are

$$y - \frac{aE_1}{M} \left(\frac{N}{2} + \frac{a(E_1 - E_2)}{2\beta} \right) + \delta(w_1 + E_1) - \theta(w_1 + E_1)^2, \tag{f1}$$

$$w_1 + E_1 = \alpha - \beta \left(\frac{N}{2} + \frac{a(E_1 - E_2)}{2\beta} \right) + E_1 \tag{f2}$$

Substituting (f2) into (f1), setting the derivative with respect to E_1 equal to zero, and then solving for E_1 yields the reaction function in (5).

response function), which gives the optimal E_1 conditional on E_2 . This function is given by

$$E_1 = \frac{\Lambda}{\Phi} + \frac{\Omega}{\Phi} E_2, \quad (5)$$

where

$$\Omega = \frac{1}{2\beta M} - \frac{\theta}{2} \left(\frac{1}{a} - \frac{1}{2} \right) \quad (6)$$

$$\Phi = \frac{1}{\beta M} + 2\theta \left(\frac{1}{a} - \frac{1}{2} \right)^2 \quad (7)$$

$$\Lambda = -\frac{N}{2M} + \left(\frac{1}{a} - \frac{1}{2} \right) Q, \quad (8)$$

with $Q = \delta - 2\theta(\alpha - \beta N/2) > 0$.¹² Observe that district 2's reaction function is given by interchanging E_1 and E_2 in (5). Note also that if the two districts had different characteristics (perhaps different rich populations (M_1 and M_2) or different wage-function intercepts (α_1 and α_2)), the positions of their reaction functions would differ. This property is reflected in the setup of the subsequent empirical model.

Recalling that $a < 1$, Φ is positive, while the signs of Ω and Λ are ambiguous. As a result, the reaction function's slope (Ω/Φ) can be either positive or negative, with spending levels being either strategic complements or substitutes. A key question is how the residency parameter a affects this slope. Inspection of (6) and (7) shows that $\partial\Omega/\partial a > 0$ and $\partial\Phi/\partial a < 0$, and these inequalities in turn imply that the reaction function's slope, equal to Ω/Φ , is increasing in a when $\Omega > 0$, or when the slope is positive.¹³ Therefore, when districts 1 and 2 interact positively, their interaction is stronger the looser is the welfare residency requirement (that is, the larger is a). When interaction is negative, however, a 's effect on the slope is ambiguous.

A further question is how the larger a after the reform affects the equilibrium level of welfare benefits. Given symmetry, this level is found by setting $E_1 = E_2$ in (5) and solving for the common value E^* . This solution, which corresponds to the intersection of the two districts' reaction functions, yields

$$E^* = \frac{\Lambda}{\Phi - \Omega} = \frac{\Lambda}{\frac{1}{2\beta M} + 2\theta \left(\frac{1}{a} - \frac{1}{2} \right)}. \quad (9)$$

It is easily shown that E^* in (9) is decreasing in a , so that the equilibrium level of welfare benefits is lower the looser is the residency requirement. This conclusion

¹² This expression equals the marginal utility of I_1 when E_1 equals zero and workers are equally divided, a positive quantity.

¹³ This conclusion follows because the slope's derivative has the sign of $\Phi \partial\Omega/\partial a - \Omega \partial\Phi/\partial a$, which is positive when $\Omega > 0$ given (6) and (7).

is intuitively sensible given that a higher a raises the tax burden from welfare payments, reducing the equilibrium generosity of property owners after the reform.

The predictions of the model change if one of its key assumptions is altered. Suppose that, instead of caring about the poor's statutory income (thus focusing on E_1), property owners care about effective income, focusing on aE_1 . Then, no new analysis is needed, and the reaction function is found by simply substituting aE_i , $i = 1, 2$, in place of E_i in Brueckner's (2000) version of the function, yielding $aE_1 = \eta + \phi aE_2$, where η and ϕ are constants. Dividing by a yields $E_1 = \eta/a + \phi E_2$. The reaction function's slope ϕ is thus independent of a , in contrast to the previous case. However, the analog to (9) sets aE^* equal to a constant, so that E^* is inversely related to a , as under the previous assumption.

Summarizing yields

Proposition 1

- (1) *If property owners care about the statutory income of the poor, then loosening of the welfare residency requirement under the reform strengthens interaction between districts when interaction is positive. The looser requirement also reduces the equilibrium level of welfare benefits, regardless of the direction of interaction.*
- (2) *If property owners instead care about the poor's effective income, then the reform has no effect on the strength of interaction but again reduces the equilibrium level of welfare benefits.*

Thus, while the equilibrium benefit level falls with the reform under both assumptions, the reform's effect on the strength of strategic interaction depends on whether property owners care about statutory or effective income, with no effect present in the latter case. The estimating equations introduced below are designed to embrace both possibilities.

The model can be extended in several ways.¹⁴ The appendix explores the case where, instead of being fixed, the income of property owners consists of profit from the production process in which poor workers are employed. In this case, which is discussed further below in interpreting the regression results, strategic interaction is stronger than in the existing model.

Another extension involves the model's portrayal of forces that retard relocation across districts. In the current setup, migration is equilibrated through changes in wages, with the wage falling via decreasing returns as workers move into a district, reducing the incentive to relocate. A related inhibiting force would arise through migration costs, although adding such costs to the model is ruled out by its static

¹⁴ As in previous models of welfare competition, the present framework does not recognize the possible coexistence of tax competition between districts. However, since the tax base supporting welfare spending was mostly immobile (with land playing a large role), the scope for tax competition during this period would appear to have been limited.

nature. However, a very similar relocation friction arises through idiosyncratic individual benefits from living in a district. To add them, the wage would be exogenously fixed at some common value w and a random term ϵ_i representing location benefits would be added to income, yielding $w + aE_i + \epsilon_i$ as the utility from living in district $i = 1, 2$. Then, assuming that random terms have a type-1 extreme value distribution with dispersion parameter σ , the probability that a worker chooses district 1 is given by the multinomial logit expression

$$\frac{\exp[(w + aE_1)/\sigma]}{\exp[(w + aE_1)/\sigma] + \exp[(w + aE_2)/\sigma]} = \frac{n_1}{N}, \quad (10)$$

where the equality sets the share n_1/N of workers living in district 1 equal to the logit probability. Differentiating (10) yields $\partial n_1/\partial E_1 > 0$, showing that a higher welfare benefit attracts more workers, as happens in the existing model (recall (2)). Also from (2), the magnitude of $\partial n_1/\partial E_1$ in the existing model is smaller the stronger is the inhibiting force from decreasing returns (measured by β). In the logit version of the model, the strength of the inhibiting force depends on the dispersion parameter σ , and calculation shows that $\partial n_1/\partial E_1$ is smaller the greater is σ . Thus, greater dispersion in locational benefits limits worker relocation in response to an increase in E_1 . The strength of the inhibiting force also affects the reaction function's slope. The β derivative of the slope expression Φ/Ω in (5) is negative, indicating that a stronger inhibiting force weakens strategic interaction when interaction is positive. However, although one might expect an analogous negative σ effect in the logit version of the model, verification is not feasible analytically since the function then has no closed-form solution. Nevertheless, the fact that the reaction-function slope can depend on migration frictions from idiosyncratic locational preferences is useful in interpreting the subsequent regression results.¹⁵

4 Empirical model and identification

Drawing on the theoretical model just presented, this section develops an empirical framework for estimating district reaction functions. Let e_0 and e_1 be the vectors of average district expenditures per pauper in the pre- and post-reform years (1857–60 and 1861–64), with each vector having dimension equal to the number of districts,

¹⁵ The strategic-interaction literature includes models of tax (and welfare) competition and yardstick competition, as well as models of interaction due to expenditure spillovers (see Brueckner (2003)). While the different sources of interaction can be hard to distinguish empirically, there appears to be no reason to think that yardstick competition or expenditure spillovers are relevant in the Poor Law environment. Yardstick competition arises in a public-service production context, where residents check to see whether the combination of taxes and public-service levels is more favorable in nearby jurisdictions (possibly due to greater efficiencies) when evaluating their own public servants. However, with poor relief not involving production of an explicit public service consumed by taxpayers, there would appear to be little scope for the inter-district comparisons underlying yardstick competition. Expenditure spillovers arise when residents consume the public services of neighboring jurisdictions (parks, museums, etc.), but poor relief generates no such spillovers.

n . Let X_0 and X_1 be the $n \times k$ matrices of k district characteristics during these periods. Let W denote the weight matrix, whose rows assign weights to the “neighbors” of a district, indicating their importance in influencing its own spending decisions. For example, the j th element of row i , denoted w_{ij} , could be set equal to the inverse of the distance between districts i and j ($w_{ij} = 1/d_{ij}$), indicating that districts near i have a greater influence than those farther away. W ’s diagonal elements are set at zero, so that a district does not influence itself.¹⁶ Then, the empirical reaction functions for the two periods, written in matrix form, are given by

$$e_0 = \gamma_0 We_0 + X_0 \lambda + u_0 \tag{11}$$

$$e_1 = \gamma_1 We_1 + \mu_1 + X_1 \lambda + u_1. \tag{12}$$

In (11) and (12), u_0 and u_1 are error vectors, and μ_1 is a column vector of dimension n with identical elements equal to the scalar μ_1 , which serves as an intercept shifter for period 1. As usual in the literature, the specification in (11, 12) assumes that district characteristics (X_0 and X_1) affect the level of the reaction function but not its slope, which is common to all districts within a given period. The period-specific slopes (the scalars γ_0 and γ_1) are allowed to differ, as suggested by the theoretical model, but X ’s coefficient vector λ is common across the periods. To better grasp the meaning of (11) and (12), observe that district i ’s row in the matrix equation (12) sets e_{1i} equal to γ_1 times the inner product of the i th row of W and the e_1 vector (the sum of the weighted e values for i ’s neighbors) plus the inner product of the i th row of X_1 (i ’s characteristics) and the coefficient vector λ (μ_1 and $u_{1,i}$ are then added).

While (11) and (12) could be estimated separately for the periods, joint estimation is required to impose the common- λ requirement. To do so, note that (12) can be written as

$$e_1 = \gamma_0 We_1 + (\gamma_1 - \gamma_0) We_1 + \mu_1 + X_1 \lambda + u_1. \tag{13}$$

Then, (11) and (12) together can be written as

$$\begin{bmatrix} e_0 \\ e_1 \end{bmatrix} = \gamma_0 \begin{bmatrix} We_0 \\ We_1 \end{bmatrix} + (\gamma_1 - \gamma_0) \begin{bmatrix} 0 \\ We_1 \end{bmatrix} + \begin{bmatrix} 0 \\ \mu_1 \end{bmatrix} + \begin{bmatrix} X_0 \\ X_1 \end{bmatrix} \lambda + \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}, \tag{14}$$

where the bracketed elements are stacked vectors or matrices. Let the stacked expenditure-per-pauper vector be denoted e and the stacked characteristics matrices and error vectors be denoted X and u . Furthermore, let

$$\begin{bmatrix} We_0 \\ We_1 \end{bmatrix} \equiv \begin{bmatrix} e_{0_lag} \\ e_{1_lag} \end{bmatrix} = e_lag, \tag{15}$$

where “_lag” denotes the spatial lag generated by W and where e_lag is a stacked vector. Then, (14) can be written as

¹⁶ Weight matrices are typically normalized in some fashion, as discussed below.

$$e = \gamma_0 e_lag + (\gamma_1 - \gamma_0) post * e_lag + \mu_1 post + X\lambda + u, \tag{16}$$

where *post* is a (scalar) dummy variable equal to 1 for post-reform observations and zero otherwise and where μ_1 is now a column vector of dimension $2n$. Since Proposition 1(i) predicts stronger interaction in period 1 than in period 0, $\gamma_1 - \gamma_0$ would then be positive, implying a positive coefficient for *post* * *e_lag*.¹⁷ Proposition 1(ii), however, predicts $\gamma_1 = \gamma_0$ and thus a zero *post* * *e_lag* coefficient.

To test the hypothesis that the reform reduces the equilibrium level of welfare spending, the proper approach is to estimate a regression that includes the post variable and the district characteristics on the right-hand side, without the appearance of *e_lag* and *post* * *e_lag*. However, equilibrium spending will depend on the characteristics of the district’s neighbors as well as its own characteristics, given that these characteristics affect the positions of both the own and neighbor reaction functions. The regression must then include the weighted sum of neighbor characteristics, equal to *WX* and denoted *X_lag*, as well as *X*, thus taking the following form:

$$e = \phi post + X\rho + X_lag \kappa + \epsilon. \tag{17}$$

Because a district’s own spending is jointly determined with neighbor spending via strategic interaction, the *e_lag* terms on the right-hand side of the reaction function in (16) are endogenous, rendering OLS estimation potentially problematic. This endogeneity, which causes *e_lag* to be correlated with the error vector *u* in (16), can be seen in the theoretical model, as follows. Let additive error terms ϵ_1 and ϵ_2 be introduced into the theoretical reaction functions of districts 1 and 2 ((5) and its counterpart for district 2), reflecting the presence of unmeasured district characteristics. Since the equilibrium spending level in (9) corresponds to the intersection of these reaction functions, that spending level will depend on both ϵ_1 and ϵ_2 . With a neighbor district’s equilibrium spending then depending on both error terms, it will be correlated with the error term in each reaction function (with ϵ_1 in 1’s function and with ϵ_2 in 2’s function), ruling out the use of OLS.

Most of the previous literature on strategic interaction addresses this problem by using instrumental variables, with the most common approach involving the use of neighbor characteristics as instruments for neighbor spending. This traditional approach is followed in the present paper, using *X_lag* as an instrument for *e_lag*.¹⁸

¹⁷ This prediction is conditional on interaction being positive. Observe also that a differencing approach could be used for joint estimation of (11) and (12). Subtracting (11) from (12) yields

$$\begin{aligned} e_1 - e_0 &= \gamma_1 We_1 - \gamma_0 We_0 + \mu_1 + (X_1 - X_0)\lambda + u_1 - u_0 \\ &= \gamma_0 We_1 - \gamma_0 We_0 + (\gamma_1 - \gamma_0)We_1 + \mu_1 + (X_1 - X_0)\lambda + u_1 - u_0 \\ \Delta e &= \gamma_0 W\Delta e + (\gamma_1 - \gamma_0)We_1 + \mu_1 + \Delta X\lambda + \Delta u, \end{aligned} \tag{f3}$$

where Δ indicates the difference between pre- and post-reform values. This equation would be estimated by 2SLS using *WΔX_lag* and *WX₀* as instruments for *WΔe* and *We₀*. This approach, however, was not successful.

¹⁸ As shown by Kelejian and Prucha (1998), the IV approach not only handles endogeneity of the spatial-lag variables but also addresses the bias and inconsistency arising from any spatial autocorrelation in the error terms.

Since own characteristics (X) help determine own spending (e), neighbor characteristics (X_{lag}) are naturally taken as determinants of neighbor spending (e_{lag}). The regression in (16), however, contains another endogenous variable, $post * e_{lag}$, and following the usual approach for instrumenting interaction terms, this variable is instrumented by $post * X_{lag}$.¹⁹

While the use of neighbor characteristics as instruments has a long tradition in spatial econometrics, as evidenced in the foundational methodological study of Kelejian and Prucha (1998), the current state-of-the-art in estimating reaction functions eschews this approach. This change has been prompted by arguments of Gibbons and Overman (2012) and others asserting that neighbor characteristics may not be excludable as covariates in a properly specified reaction function, thus rendering them improper instruments. Under this view, in choosing its own expenditures, a district may take into account both the expenditures (e_{lag}) of neighboring districts *and* their characteristics (X_{lag}), requiring a search for instruments other than X_{lag} .²⁰ Responding to such criticisms of the traditional approach, Dahlberg and Edmark (2008), Baskaran (2014), Lyytikäinen (2012), Agrawal (2015), and Parchët (2019) cleverly exploit institutional features of their empirical settings to generate alternate instruments with stronger excludability justifications.²¹ Since the 19th-century empirical setting appears not to offer such opportunities, the analysis falls back on the traditional approach, using neighbor characteristics as instrument in a 2SLS (two-stage-least-squares) regression based on (16). Regression diagnostics (the results of the Sargan overidentification test) suggest, however, that neighbor characteristics are valid instruments in the current setting.

¹⁹ Given the resulting complexity, higher powers of W times X_{lag} and $post * X_{lag}$ are not used as instruments, in contrast to the suggestion of Kelejian and Prucha (1998). Thus, the characteristics of the neighbors of the neighbors of a district (given by W^2X) do not appear as instruments.

²⁰ This claim can be checked in the theoretical model of Sect. 3. Suppose, as suggested above, that both the rich population (M) and the wage-function intercept (α) are district specific. Then, M in footnote 11 is replaced by M_1 and α by α_1 , with no appearance of either M_2 or α_2 . The reaction function generated by differentiating (1) thus does not depend on these neighbor characteristics, contrary to the claim. However, if the wage-function slope (β) is district specific, then both β_1 and β_2 will appear in both districts' reaction functions, so that neighbor characteristics along with own-characteristics appear in each district's function. The upshot is that neighbor characteristics may or may not appear in the reaction functions, depending on the model setup.

²¹ Parchët (2019), for example, estimates a model of strategic interaction in the setting of municipal income tax rates in Switzerland, where cantonal (state-level) income taxes are added to the municipal tax to get the combined income-tax rate and thus the overall tax burden. A municipality's own income-tax rate is set conditional on its own state tax and on the combined tax of neighboring municipalities. The natural assumption that a municipality cares about the combined tax rate of its neighbors, rather than the separate municipal and state components, allows the neighbor's state tax rate to be used as an instrument for its combined tax rate.

5 Data

The empirical analysis draws on newly collected, comprehensive, and spatially disaggregate Poor Law data taken from the annual reports of the Poor Law Board.²² Covering the entirety of England and Wales, these reports provide detailed information on welfare expenditures by program, information on the numbers and demographic composition of paupers relieved through these programs, data on Poor Law revenues, and, in some years, the value of local property subject to taxation. Information is reported at the level of Poor Law Unions, which, while technically distinct from registration districts (the main administrative units at the time), map roughly to the latter units. Thus, in order to account for changing boundaries over time, and to allow compatibility with data reported at the registration district-level (mainly census data), a panel dataset is constructed consisting of 536 standardized registration districts covering all of England and Wales over the period 1857–1864.²³ Figure 1 shows a district map, with darker shades indicating higher average welfare-benefit levels over the pre-reform period (1857–1860). The darkest districts have levels above £ 7.91 and lightest have levels below £ 6.05.

The analysis begins in 1857 when data on pauperage becomes available, allowing welfare expenditure per recipient to be computed. In order to avoid the confounding effect of a second welfare eligibility change in 1865 (see footnote 9), the analysis ends in 1864. The timing of the new residency requirement, which came into effect in 1861, allows construction of balanced pre-reform (1857–1860) and post-reform (1861–1864) periods. The annual data are averaged within each of these periods, allowing the strength of interaction in welfare spending to be estimated in each of two 4-year periods and compared between them. Note that collapsing the data in this way allows for a simple before-and-after empirical structure, while using the eight individual years of underlying annual data would yield greater complexity with little apparent benefit (for example, $4 \times 4 = 16$ possible interaction comparisons between pre- and post-reform years).

The dependent variable in the analysis is the district's average annual expenditure per pauper over the period, denoted **epp**, which is the empirical analog to the *E* and *e* variables from Sects. 3 and 4. The variable is computed by dividing total welfare spending in pounds sterling (including both indoor and outdoor relief) by a count of the number of recipients (paupers), thus mixing the two types of relief despite a difference in costliness. Table 1, which presents summary statistics for the sample, shows a mean **epp** value of £ 7.06 across both subperiods, with the post-reform mean of £ 6.96 being slightly lower than the pre-reform mean of £ 7.17. The range of payments across districts is wide, with standard deviations above 2, a pattern that Clark and Page (2019) believe shows wide regional differences in generosity toward the poor. Observe that while the pre-post difference in the **epp** means is predicted in Proposition 1, a proper test requires a full regression, as seen below. Note also that a

²² The data were collected by my UCI colleague Vellore Arthi and her coauthors Brian Beach and W. Walker Hanlon (see Arthi et al., 2017a).

²³ Three districts are omitted because of missing values for key variables.

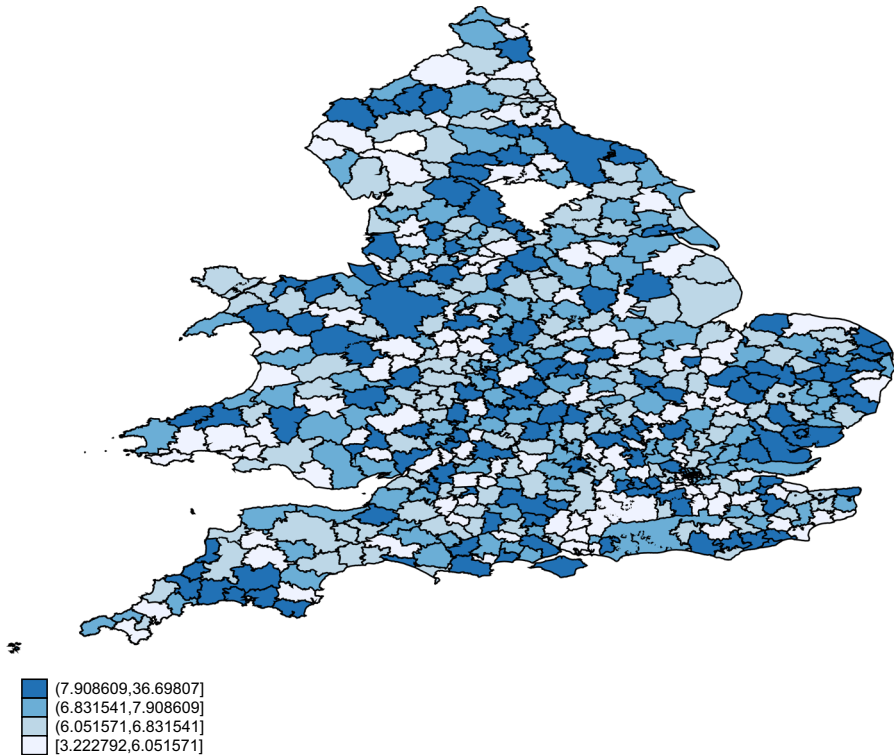


Fig. 1 District benefit levels, England and Wales

limitation of the **epp** variable is that the pauper count in the denominator, which is an average of those receiving relief in January and July of the year, does not account for differences in the duration of relief over the year across paupers.

The key explanatory variables are **epp_lag**, the average over the subperiod of neighboring districts' weighted annual expenditure per pauper, and the interaction of this variable with a **post** indicator for the 1861-1864 post-reform period, denoted **post*epp_lag**. The coefficient of **epp_lag** gives the baseline (pre-reform) extent of interaction, while the coefficient of **post*epp_lag** measures the change in the extent of interaction due to the relaxation of residency requirements (see (16)).

As explained earlier, **epp_lag** is equal to the weight matrix W times **epp**, with **epp** viewed as a vector (corresponding to e_{lag} from Sect. 4). Results using four different weight matrices are presented below. The first is a contiguity matrix, which assigns a weight of 1 to districts contiguous to the given district and zero otherwise. The second matrix uses inverse distance weights, as discussed in Sect. 4, but these weights are truncated (being set to zero) beyond a distance of 30 km. This matrix is

Table 1 Summary statistics

Variables	Obs.	Mean	Std. Dev.	Min	Max
epp	1072	7.06	2.04	3.22	36.70
—epp (post = 0)	536	7.17	2.32	3.22	36.70
—epp (post = 1)	536	6.96	1.70	3.66	21.76
epp_lag	1072	1.46	2.36	0.039	18.96
—epp_lag (post = 0)	536	1.51	2.48	0.039	18.96
—epp_lag (post = 1)	536	1.40	2.23	0.040	17.49
population	1072	37,429	41,772	2359	341,410
—population (post = 0)	536	36,518	40,311	2471	319,506
—population (post = 1)	536	38,340	43,202	2359	341,410
big_city	1072	0.05	0.22	0	1
cotton_district	1072	0.04	0.21	0	1
rateable_value	1072	126,578	131,180	8918	1,115,551
pop_shr_under_5	1072	0.131	0.009	0.098	0.163
pop_shr_55_up	1072	0.114	0.019	0.060	0.167
emp_shr_wc	1072	0.040	0.021	0.014	0.132
emp_shr_nontrade	1072	0.111	0.032	0.056	0.303
emp_shr_manuf	1072	0.122	0.082	0.021	0.425
emp_shr_trans	1072	0.020	0.011	0.004	0.085
emp_shr_other	1072	0.198	0.068	0.044	0.362
netmig_rate	536	− 0.051	0.126	− 0.296	0.548
inmig_rate	536	− 0.082	0.066	− 0.296	0

Computation of epp_lag uses the Invdpop75 weight matrix

denoted *Invdist30*, while the *Invdist75* matrix truncates the inverse distance weights at zero beyond a distance of 75 km, rather than 30 km. These relatively short distances were those over which the bulk of inter-district migration in this period typically took place (Arthi et al., 2020). The last two weight matrices, denoted *Invdpop30* and *Invdpop75*, use truncated weights that depend on both neighbors' population and distance. The matrix element w_{ij} equals P_j/d_{ij} when d_{ij} is less than either 30 or 75 km for the two matrices and zero otherwise, where P_j denotes district j 's pre-reform population. Thus, another district j receives a larger weight when it is more populous and a smaller weight when it is farther from district i .

In estimating the reaction functions, the weight matrices are normalized to produce coefficients of convenient sizes. "Spectral" normalization is used, which rescales all the matrix elements by a common factor that generates a largest eigenvalue of 1 for the matrix.²⁴ The scaling of the matrix affects the size of the estimated reaction-function slope coefficient (γ_0 in (15)) but not its sign or statistical significance. As a result, the estimates based on the normalized weight matrices can

²⁴ Row normalization is commonly used, but this method is dispreferred for distance-related weight matrices, being more appropriate for contiguity weights.

show whether or not strategic interaction is occurring among districts (whether the estimated γ_0 is significantly different from zero). The quantitative degree of interaction can then be gauged by rescaling the estimated γ_0 to undo the effects of the normalization. Given spectral normalization of the weight matrix, the actual mean value of **epp_lag** shown in Table 1 (which is based on the *Invdpop75* matrix) is not meaningful. Note, however, that the subperiod **epp_lag** mean decreases between pre- and post-reform periods, mirroring the decrease in **epp**.

Additional district characteristics are drawn from decennial censuses, which provide information on demographics and occupational structure, and from the annual reports of the Registrar General, which provide information on vital events. District population is computed through interpolation by combining decennial census counts with annual information on intercensal births and deaths. It is measured annually, with average values (as in the expenditures data) being computed for each of the pre- and post-reform periods.²⁵ This variable is denoted **pop**, and its overall mean is 37,429, with the subperiod means showing slight population growth between the subperiods (**pop** is rescaled to units of 10,000 in the regressions).

In addition to population, the regressions include a number of other district characteristics. They are presented in Table 1 and include a dummy variable **bigcity**, which indicates that a district lies in one of the major metropolitan areas of the time, London, Manchester, Liverpool or Greater Leeds (mean = 0.05); a dummy variable **cottondistrict**, which indicates a district where employment was concentrated in cotton processing, one of the most important industries of the time (mean = 0.04); a variable **rateable_value**, a tax-base measure that gives the value in pounds of district property subject to taxation (mean = \$ 126,578, but rescaled to units of 100,000 in the regressions);²⁶ and a series of variables whose names contain “shr” (indicating share), which capture district demographic and occupational characteristics. The population-share variables are **pop_shr_under_5** and **pop_shr_55_up**, which give the district population shares under 5 years and at least 55 years of age, respectively (means = 0.131 and 0.114; note that ages 5-54 are the omitted age category). These variables are meant to capture the effect of typically vulnerable populations on local welfare spending.

The employment-share variables are denoted **emp_shr_xx**, where **xx** = **wc** (white collar, mean = 0.040), **nontrade** (nontradeable goods, mean = .111), **manuf** (manufacturing, mean = 0.122), **trans** (transportation, mean = 0.020), and **other** (other types of non-agricultural employment, mean = 0.198), with the share in agricultural employment acting as the omitted category. These variables are meant to capture the impact of local labor markets on welfare spending decisions. Aside from the population variable, all of these district characteristics are time invariant, being tabulated using data from the year 1851, the most recent census year preceding the sample

²⁵ Population is interpolated via the Das Gupta method (a components-of-change approach), using the Census years of 1851, 1861, and 1871, along with intervening mortality and fertility statistics. See Arthi et al. (2017b) for details of the procedure.

²⁶ Because property values were reported only sporadically, a value for 1851 is not available, with rateable values instead taken from the 1852 report of the Poor Law Board.

period. Data from this year yield a clean, pre-intervention snapshot of key district features. Although it is technically feasible to generate separate post-reform values from 1861 sources, this approach is undesirable since district characteristics may have changed endogenously in response to the reform. In any case, the characteristics used in the analysis are fairly stable across the 1851 and 1861 censuses.

Finally, the regressions include alternate sets of fixed effects. County fixed effects indicate which of the 55 counties in England and Wales the district belongs to, while division fixed effects indicate which of 12 divisions (a broader spatial unit) contains the district. These fixed effects help address another threat to identification: the existence of common unobserved factors (common shocks) that cause welfare spending to move in step across nearby districts, potentially creating a false impression of positive strategic interaction when none exists (in this case, the error terms are spatially autocorrelated). The IV approach addresses this problem while also remedying simultaneity bias, but using fixed effects to control for common shocks further helps to limit any resulting upward bias.²⁷

Even though unspecified inter-district error correlation can be remedied via the IV approach, the structure of the data also implies the presence of intra-district error correlation. In particular, elements of the pre- and post-reform error vectors (u_0 and u_1) that pertain to the same district will be correlated due to common unobservables. The remedy is to cluster coefficient standard errors by district. Thus, the regressions combine either division or county fixed effects with clustering by district.

6 Empirical findings

6.1 The reform's effect on equilibrium spending

One prediction of the model is that equilibrium welfare spending will be lower following the loosening of residency requirement. This hypothesis is tested by estimating Eq. (17) with either division or county fixed effects and standard errors clustered at the district level. In each case, the regression includes $X_{lag} = WX$ (spatial lags of the district characteristics), but their coefficients are not shown. Since the remaining coefficients are not very sensitive to the choice of W , results using *Invdpop75* weights are reported.

The results are shown in Table 2, and as can be seen, the **post** coefficient is significantly negative in each regression, with its magnitude indicating that welfare expenditures per pauper fall by about 1/4 of a pound following the reform, equivalent to roughly 30 days' bare bones subsistence costs in mid-19th century England

²⁷ Use of district fixed effects is feasible in principle, but with only two time periods in the sample, this approach removes a great deal of the variation in the data. Thus, use of district-level controls is preferred, along with higher-geography fixed effects. Note that use of district fixed effects is equivalent to estimating the first-difference model described in footnote 17, which was not successful.

Table 2 Determinants of equilibrium spending

Variables	(1)	(2)	(3)	(4)
post	- 0.220** (0.0511)	- 0.218** (0.0498)	- 0.213** (0.0521)	- 0.213** (0.0532)
pop	- 0.0228 (0.0207)	- 0.0360 (0.0219)	-	-
bigcity	0.519 (1.261)	0.250 (1.357)	-	-
cottondist	0.108 (0.281)	- 0.0612 (0.396)	-	-
rateable_value	0.0769 (0.113)	0.0811 (0.119)	-	-
pop_shr_under_5	- 40.61** (14.32)	- 37.44** (13.85)	- 29.88** (10.17)	- 37.75** (10.18)
pop_shr_55_up	- 14.47 (7.702)	- 8.630 (8.421)	-	-
emp_shr_wc	- 17.30* (7.982)	- 7.985 (7.889)	-	-
emp_shr_nontrade	9.606 (5.099)	9.307 (5.507)	6.024 (3.438)	4.894 (3.440)
emp_shr_manuf	- 1.385 (2.123)	- 1.391 (2.631)	- 3.104** (1.189)	- 3.770* (1.528)
emp_shr_trans	11.12 (8.908)	4.279 (9.018)	-	-
emp_shr_other	3.701 (2.779)	3.414 (3.213)	-	-
Constant	13.07** (3.162)	9.382** (2.978)	10.63** (1.448)	9.347** (1.454)
Fixed effects	<i>Division</i>	<i>County</i>	<i>Division</i>	<i>County</i>
Clustered standard errors	<i>yes</i>	<i>yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	1072	1072	1072	1072
R ²	0.344	0.403	0.321	0.377

Robust standard errors in parentheses ** $p < 0.01$, * $p < 0.05$. Additional covariates are spatial lags of district characteristics ($X_{lag} = WX$), where $Invdpop75$ weights are used

(Allen, 2015; Humphries and Weisdorf, 2019). This finding supports the predictions of both versions of the theoretical model.²⁸

²⁸ If total welfare spending were left unchanged in each district following the reform, then, with a greater number of poor being eligible even in the absence of migration, spending per pauper would fall in a mechanical fashion. However, this scenario, which would also generate the results seen in Table 2, does not conform to the district behavior portrayed in the model, where total spending responds to the increase in eligibility.

In columns 1 and 2, where the full list of covariates is used, just one of the district characteristics has a statistically significant effect on the equilibrium benefit levels. Spending is lower in districts with a large share of young children, with the **pop_shr_under_5** coefficient significantly negative in both regressions. The explanation for this young-children effect is not entirely clear, but their presence may indicate a younger overall district age structure (including parents), resulting in lower demand for welfare support. Alternatively, with children having lower per capita subsistence needs, welfare payments to their families could then be lower.

Collinearity of some explanatory variables may block significant effects from emerging in columns 1 and 2, and restriction of the set of covariates, as seen in columns 3 and 4, yields one other significant determinant of welfare benefits: share of employment in manufacturing. The **emp_shr_manuf** coefficients are significantly negative along with the coefficient of **pop_shr_under_5**, presumably reflecting a lower need for welfare spending in districts with substantial well-paid manufacturing employment. While the coefficients of the non-tradeable employment share are insignificant, this variable is included because of its later role as a reaction-function shifter.²⁹ Note finally that the qualitative results in Table 2 are mostly independent of whether division or county fixed effects are used.

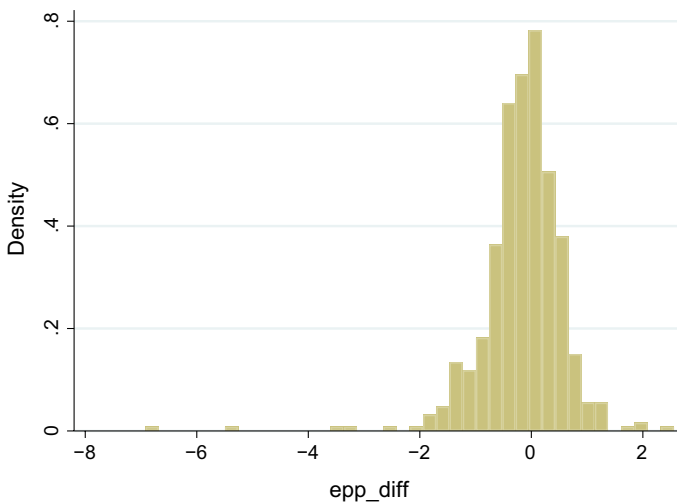


Fig. 2 Histogram of epp changes

²⁹ The individual effects of the variables with insignificant coefficients in columns 1 and 2 can also be gauged in separate regressions where each one appears by itself along with **post**. In these regressions, the coefficients of **emp_shr_wc**, **emp_shr_nontrade**, and **emp_shr_other** are significantly positive, although these positive effects disappear in regressions that include more covariates. While **pop_shr_55_up** has a significantly positive coefficient when it appears by itself, the coefficient becomes insignificant in the presence of **pop_shr_under_5**. This variable, which thus appears inessential, is not included in subsequent regressions because some of the reaction function estimations conform better to expectations without it.

The origin of the negative **post** coefficients in Table 2 can be seen in Fig. 2, which shows a histogram for the variable **epp_diff**, equal to the difference between the post- and pre-reform **epp** values. In reading the histogram, recall that **epp** is measured in pounds and that the mean pre- and post-reform values bracket £ 7. As the histogram shows, **epp_diff** is more often negative than positive, with some values substantially below zero (two even larger negative outliers were dropped to improve the readability of the graph). The histogram is only a partial explanation for the negative **post** coefficients, however, given that the regression estimates hold other factors constant.

6.2 Reaction functions

Before formally testing for strategic interaction, it is useful to examine whether spatial dependence is present in this setting by performing Moran's I test. In the present context, this procedure tests for spatial correlation of the weighted residuals from a regression of **epp** on a set of covariates, done separately for the pre- and post-reform periods. When **epp** is regressed on a constant term, the test strongly rejects the absence of spatial dependence in both periods, using *Invdpop75* weights. However, when the regression uses the covariates in Table 2, the test has an (insignificant) *p*-value of 0.13 in the pre-reform period, although absence of spatial dependence is strongly rejected in the post-reform period. With some evidence of spatial dependence in hand, the analysis turns to estimation of reaction functions.

Table 3 shows reaction functions estimated using contiguity weights along with *Invdist30* and *Invdpop30* weights, using mostly the same set of district characteristics as in Table 2, division fixed effects, and clustered standard errors (results are similar with county fixed effects). Both the **epp_lag** and **post * epp_lag** coefficients are statistically insignificant in all three regressions. Thus, the somewhat surprising lesson of Table 3 is that, when the neighboring districts that receive positive weights are close by, whether contiguous or within 30 km, then strategic interaction appears to be absent. Note that the regression diagnostics shown at the bottom of the table are favorable. The Cragg-Donald F statistics are large, allaying any concerns about weak instruments, and the Sargan overidentification test's *p*-values are well above the critical 5% value, suggesting that the instruments are valid.

Table 4 shows reaction function estimates when the range of nonzero weights is extended to 75 km. The regressions use *Invdpop75* weights, where district *j*'s weight is P_j/d_{ij} when $d_{ij} < 75$ km and zero otherwise (*Invdist75* weights, equal to $1/d_{ij}$, yield similar results). Columns 1–3 show regressions without clustered standard errors, and among these unclustered regressions, column 1 shows an OLS regression for comparison purposes, while columns 2 and 3 are 2SLS regressions with division and county fixed effects, respectively. Clustering of standard errors is added in columns 4 and 5.

As can be seen, the results of Table 3 are mostly overturned in Table 4. Looking first at the unclustered regressions in columns 1–3, the **epp_lag** coefficients are significantly positive. As usual, clustering increases the estimated standard error of the **epp_lag** coefficient, as can be seen by comparing columns 2 and 4 and columns 3

Table 3 Reaction functions with contiguity, Invdist30, and Invpop30 weights

Variables	(1) Contiguity	(2) Invdist30	(3) Invpop30
epp_lag	- 0.00361 (0.0310)	0.155 (0.118)	0.170 (0.0911)
epp_lag_post	0.0201 (0.0192)	- 0.0442 (0.0336)	- 0.0283 (0.0298)
post	- 0.323* (0.127)	- 0.154* (0.0613)	- 0.175** (0.0553)
pop	- 0.0195 (0.0213)	- 0.0131 (0.0199)	- 0.0105 (0.0192)
bigcity	1.081 (1.114)	0.832 (1.274)	0.732 (1.255)
cottondist	0.200 (0.320)	0.170 (0.318)	0.123 (0.320)
rateable_value	0.103 (0.122)	0.0930 (0.119)	0.0926 (0.119)
pop_shr_under_5	- 24.59* (12.31)	- 24.76* (12.29)	- 25.34* (12.25)
emp_shr_wc	- 11.41 (7.436)	- 10.95 (7.189)	- 11.37 (7.169)
emp_shr_nontrade	13.34* (5.294)	12.68* (5.576)	12.55* (5.415)
emp_shr_manuf	- 2.639 (2.225)	- 2.854 (2.223)	- 2.859 (2.219)
emp_shr_trans	6.478 (10.27)	4.601 (9.803)	5.366 (10.08)
emp_shr_other	1.333 (2.598)	1.456 (2.593)	1.594 (2.598)
Constant	9.192** (2.493)	9.123** (2.526)	9.214** (2.520)
Cragg-Donald Wald <i>F</i> stat.	319.1	7765.0	1.2e+04
Sargan overid <i>p</i> -value	0.8425	0.6502	0.1092
Fixed effects	<i>Division</i>	<i>Division</i>	<i>Division</i>
Clustered standard errors	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	1072	1072	1072

Robust standard errors in parentheses ** $p < 0.01$, * $p < 0.05$. The spatial-lag variables epp_lag and epp_lag_post are endogenous, and the 2SLS instruments are X_{lag} and $post * X_{lag}$, where X is the matrix of district characteristics variables in the regression

Table 4 Reaction functions with Invdpop75 weights

Variables	(1)	(2)	(3)	(4)	(5)
epp_lag	0.209** (0.0583)	0.221** (0.0577)	0.157* (0.0653)	0.221** (0.0851)	0.157 (0.0943)
epp_lag_post	-0.0256 (0.0444)	-0.0266 (0.0439)	-0.0332 (0.0425)	-0.0266 (0.0289)	-0.0332 (0.0285)
post	-0.153 (0.122)	-0.150 (0.120)	-0.147 (0.116)	-0.150* (0.0624)	-0.147* (0.0619)
pop	-0.00915 (0.0187)	-0.00859 (0.0184)	-0.0168 (0.0185)	-0.00859 (0.0185)	-0.0168 (0.0194)
bigcity	0.654 (0.722)	0.632 (0.714)	0.943 (0.707)	0.632 (1.283)	0.943 (1.313)
cottondist	0.0983 (0.353)	0.0927 (0.348)	0.0687 (0.365)	0.0927 (0.321)	0.0687 (0.393)
rateable_value	0.0881 (0.0665)	0.0873 (0.0657)	0.0617 (0.0650)	0.0873 (0.118)	0.0617 (0.126)
pop_shr_under_5	-27.11** (9.108)	-27.25** (9.001)	-31.50** (9.667)	-27.25* (12.21)	-31.50** (10.52)
emp_shr_wc	-10.17 (6.923)	-10.10 (6.841)	-1.725 (7.100)	-10.10 (7.160)	-1.725 (7.213)
emp_shr_nontrade	11.69** (3.805)	11.60** (3.760)	10.86** (4.031)	11.60* (5.548)	10.86 (5.917)
emp_shr_manuf	-3.216* (1.515)	-3.247* (1.497)	-2.877 (1.671)	-3.247 (2.264)	-2.877 (2.869)
emp_shr_trans	4.383 (8.654)	4.267 (8.552)	-1.617 (8.768)	4.267 (10.06)	-1.617 (10.56)
emp_shr_other	1.429 (2.110)	1.434 (2.086)	2.896 (2.337)	1.434 (2.570)	2.896 (2.753)
Constant	9.433** (1.931)	9.447** (1.908)	7.352** (2.335)	9.447** (2.522)	—
Cragg-Donald Wald <i>F</i> stat.	—	1068.1	1067.2	151.1	115.2
Sargan overid <i>p</i> -value	—	0.1341	0.2973	0.0842	0.0734
Fixed effects	<i>Division</i>	<i>Division</i>	<i>County</i>	<i>Division</i>	<i>County</i>
Clustered standard errors	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>
Observations	1072	1072	1072	1072	1072
<i>R</i> ²	0.326	—	—	—	—

Standard errors in parentheses ** $p < 0.01$, * $p < 0.05$. Column 1 is an OLS regression. The spatial-lag variables epp_lag and epp_lag_post are endogenous, and the 2SLS instruments are X_{lag} and $post * X_{lag}$, where X is the matrix of district characteristics variables in the regression

and 5. Despite this increase, the **epp_lag** coefficient in column 4 remains significant, although the coefficient becomes marginally significant in column 5 (being significant at the 10% level).³⁰ Even with this marginal significance, however, the weight of the evidence in Table 4 appears to suggest the presence of strategic interaction in the choice of welfare benefits, with spending levels per pauper being strategic complements across neighboring districts. By contrast, the **post * epp_lag** coefficients are insignificant in all of the regressions, providing no evidence of a post-reform strengthening of strategic interaction.³¹

Comparing the OLS **epp_lag** coefficient in column 1 to the larger 2SLS coefficient in column 2 (both of which rely on division fixed effects), it appears that the IV approach removes a slight downward bias in the estimated strength of strategic interaction. In addition, comparing columns 2 and 3 (or 4 and 5), the reduction in the size of **epp_lag** coefficient with county fixed effects shows that these finer fixed effects do a better job than division fixed effects of controlling for unobserved common factors, which can give a false (or elevated) impression of strategic interaction when ignored.

While the **post** coefficient is insignificant in columns 1–3 of Table 4, it becomes significantly negative with clustering in columns 4 and 5, indicating that the reaction function shifts down after the reform. As in Table 2, only a few of the district characteristics tend to have significant coefficients, indicating that these characteristics shift the reaction function. As was true in Table 2, the **pop_shr_under_5** coefficient is significantly negative in all Table 4's regressions, indicating that a high share of young children shifts the reaction function downward. The employment share of nontradables (**emp_shr_nontrade**) is also a significantly positive reaction-function shifter in columns 1–4, although the coefficient becomes marginally significant in column 5. The manufacturing employment share (**emp_shr_manuf**) has significantly negative coefficients, but only in columns 1 and 2.

As in Table 3, the diagnostic information reported at the bottom of Table 4 is acceptable. The *F* statistics for the instruments are large in all four regressions, again dispelling any concerns about weak instruments. The *p*-value for Sargan's test of overidentifying restrictions is larger than the critical 5% level in all of the 2SLS regressions, although the values in columns 4 and 5 are not substantially larger than this critical value.

The regressions in Table 4 contain a large number of instruments, corresponding to the spatially lagged values of each of the 11 district characteristics. As in Table 2, it is helpful to investigate a more parsimonious specification, where variables that appear not to be significant reaction-function shifters are dropped, cutting the number of instruments to a more modest level. Table 5 shows more-parsimonious

³⁰ Running the regression with county fixed effects and clustering required use of the "partial" option in Stata's `ivreg2`, which is used to "partial out" the county fixed effects in order to compute a well-behaved variance-covariance matrix (an error message is generated otherwise). Under this method, the county coefficients and the constant are not reported, explaining the lack of a constant in Tables 4 and 5.

³¹ When the pre- and post-reform reaction functions (11) and (12) are estimated separately, the slope coefficients are very close in magnitude, as would be expected given the insignificance of the **post * epp_lag** coefficients.

Table 5 Parsimonious reaction functions with Invdpop75 weights

Variables	(1)	(2)	(3)	(4)	(5)
epp_lag	0.227** (0.0567)	0.246** (0.0566)	0.186** (0.0625)	0.246** (0.0852)	0.186* (0.0915)
epp_lag_post	-0.0241 (0.0445)	-0.0250 (0.0441)	-0.0307 (0.0427)	-0.0250 (0.0289)	-0.0307 (0.0283)
post	-0.155 (0.122)	-0.152 (0.121)	-0.150 (0.116)	-0.152* (0.0630)	-0.150* (0.0628)
pop_shr_under_5	-32.83** (7.704)	-33.03** (7.640)	-39.69** (8.299)	-33.03** (9.904)	-39.69** (9.784)
emp_shr_nontrade	6.743** (2.431)	6.591** (2.411)	6.010* (2.482)	6.591* (3.326)	6.010 (3.385)
emp_shr_manuf	-4.173** (0.766)	-4.232** (0.760)	-4.914** (0.838)	-4.232** (1.126)	-4.914** (1.640)
Constant	10.78** (1.187)	10.81** (1.177)	9.543** (1.712)	10.81** (1.417)	-
Cragg-Donald Wald F stat.	-	1.4e+04	1057.4	1.4e+04	1.2e+04
Sargan overid p -value	-	0.8830	0.1122	0.2318	0.1093
Fixed effects	<i>Division</i>	<i>Division</i>	<i>County</i>	<i>Division</i>	<i>County</i>
Clustered standard errors	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>
Observations	1072	1072	1072	1072	1072
R^2	0.321	-	-	-	-

Standard errors in parentheses ** $p < 0.01$, * $p < 0.05$. Column 1 is an OLS regression. The spatial-lag variables epp_lag and epp_lag_post are endogenous, and the 2SLS instruments are X_{lag} and $post * X_{lag}$, where X is the matrix of district characteristics variables in the regression

regressions, using the same structure as Table 4. As can be seen, the **epp_lag** coefficients are significant in each of the regressions, including the clustered regression with county fixed effects in column 5, where marginal significance is seen in Table 4. However, as in Table 4, the **post * epp_lag** coefficient remains insignificant in all the regressions.

As for the reaction-function shifters, the coefficients are mostly the same as in Table 4. The **post** coefficients are again significantly negative in columns 4 and 5, the **pop_shr_under_5** coefficients are significantly negative in all columns, and the **emp_shr_nontrade** coefficients are significantly positive except in column 5. Now, however, the coefficients of **emp_shr_manuf** are significantly negative in all the regressions, not just in columns 1 and 2, indicating that a higher manufacturing share shifts the reaction function down. The diagnostics again show large F statistics for the instruments, and the Sargan p -values are now all well above 5%.

The results in Tables 4 and 5 establish two main patterns. First, they show qualitative evidence that districts interacted positively in their choice of welfare benefits under the New Poor Law. That is, a district would tend to raise its benefit level in response to an increase in its neighbors' benefit levels, and would tend

to reduce its benefit level in response to a neighboring reduction. This finding is intuitively sensible, even though negative interaction is possible under the model. Given that districts interact positively, however, Proposition 1-1 predicts that strategic interaction among them should strengthen as residency requirements are loosened. But the results show no significant evidence of a post-reform change in the intensity of interaction, matching the predictions of Proposition 1-2 and suggesting that property owners may have considered the effective rather than statutory income of the poor in setting welfare benefits. This alternate version of the model may thus offer a better portrayal of how welfare benefits were chosen.

Another possible explanation for lack of strengthening in strategic interaction is that the first version of the model is accurate but that the loosening of residency requirements was not large enough to be compelling to prospective migrants, relative to other considerations, such as migration costs, district wage differentials, and local unemployment rates. Moreover, even if individuals were sensitive to these changes in migration incentives, they may have been slow to respond.

Yet another explanation comes from evidence showing that Poor Law Unions exercised the right of removal for ineligible relief applicants very sparingly, perhaps making the residency requirement toothless. For instance, although migrants often made up a large share of the local adult population (up to 70% in some cities of the industrial North) only 10–15% of non-settled relief applicants in these areas were removed. Moreover, Unions were very selective in their exercise of this authority: removals mostly focused on economically undesirable migrants (single women, older workers, the permanently disabled, those in declining occupations, and those with very large families; Boyer, 1993). Such laxity in enforcement would suggest that loosening of the residency requirement may have had less effect on actual practice than envisioned by the national authorities, possibly helping to explain the lack of evidence of stronger post-reform interaction. Note, however, that this explanation and the previous one do not match up with the post-reform decline in welfare spending seen in Table 2, which is generated under both versions of the model by the greater threat of welfare migration that these explanations discount.

A final question concerns the spatial range of strategic interaction. Table 3 showed no interaction within a radius of 30 km, while Table 4 showed its emergence when the radius was extended to 75 km. If mobility were indeed low, one might have expected interaction to exist at 30 km but that inclusion of additional districts beyond 30 km (perhaps outside migration range) would have had no effect on its estimated strength. The opposite pattern may be emerging because a more-sizeable radius is needed to include those neighbors that a given district viewed as its main welfare competitors. In other words, all neighbors may not have mattered, and to ensure those that did are captured in the regression, a 75 km radius is needed. Note that, in principle, a remedy for the lack of clarity about the appropriate weight matrix would be inter-district total migration data showing origins and destinations, which would show connections between districts and could be used to specify weights. Such data are unavailable, however, for the sample period.

6.3 Quantitative effects

Finally, a quantitative interpretation of the previous results is useful. As mentioned above, spectral normalization of the weight matrices means that the **epp_lag** coefficients provide only qualitative, rather than quantitative, evidence on strategic interaction, showing whether or not it occurs without transparently revealing the magnitude of the effect. Computation of this magnitude makes use of coefficient estimates based on the un-normalized weight matrices. The **epp_lag** coefficient associated with the unnormalized *Invdpop75* matrix equals $4.56e-07$ when county fixed effects are used in the parsimonious specification of Table 5. To compute the impact of neighbor spending on a district i 's **epp** value, suppose that spending in each neighbor district j rises by Δ . Then, with *Invdpop75* weights, the value of **epp_lag** rises by $\sum_j \Delta P_j / d_{ij}$ for districts with nonzero weights. Multiplying this expression by the un-normalized **epp_lag** coefficient of $4.56e-07$ then gives the change in district i 's **epp** value, equal to $4.56e-07 \sum_j \Delta P_j / d_{ij}$. This expression differs across districts i , but the mean value can be used as a representative magnitude. Using this mean value with Δ set equal to 1, the calculation yields 0.04 as the change in a representative district's own-**epp** value. Therefore, in response to a unit increase in its neighbors' **epp** values, a district increases its own spending by 0.04, or 1/25th as much. This number becomes larger if neighboring districts tend to be closer or more populous, raising the value of the summation from above. Using the 90th percentile value (rather than the mean) of the summation, the impact on own spending from a unit increase in neighbor spending is 0.08, with spending rising 1/12th as much as neighbor spending.

The conclusion is that, even though strategic interaction appears to exist among British districts, its estimated magnitude is small. This finding could have several explanations. Most prominently, low mobility may have meant that migration of the poor across districts in response to welfare-benefit differentials was not substantial, reducing the threat of welfare migration and thus a district's attention to its neighbors' benefit levels. As was seen in the logit version of the theoretical model, low mobility arises from wide dispersion of idiosyncratic locational benefits, and although it could not be verified formally, greater dispersion may also lead to weaker strategic interaction in that model. Therefore, low mobility and weak interaction may go hand in hand, perhaps explaining the current results.³²

Looking at migration directly may shed light on whether mobility was sufficient to allow relocation to districts with generous welfare benefits. To this end, Table 6 reports regressions relating inter-district migration over the 1851-1861 period to the pre-reform **epp** level (the average over 1857-1861) along with pre-reform values of the other covariates in Table 4. In the first column, the dependent variable is the net-migration rate, denoted **netmig_rate** (which takes negative (positive) values as

³² While low mobility is likely to reduce the strength of interaction, it also may reduce the number of jurisdictions that compete, as discussed earlier in reference to the spatial range of interaction. The result can be greater market power for these competitors, which could in turn raise the strength of interaction. As a result, the ultimate connection between mobility and the strength of interaction could be unclear.

in-migration is greater than (less than) out-migration), while the dependent variable in the second column is **inmig_rate**, generated by setting the positive values of **netmig_rate** equal to zero (see Table 1 for summary statistics).³³ As can be seen in Table 6, a district's level of expenditure per pauper has no effect on either migration variable.³⁴

There are two reasons, however, why these regressions may not be informative about the extent of welfare migration. First, since welfare migrants may have represented only a small share of total migrants, benefit levels may have been overshadowed as a determinant of the overall flow. In addition, welfare migration and benefit levels may have been jointly determined, with appreciable in-migration leading districts to keep benefits low, a possibility that would bias the benefit coefficient toward zero.³⁵ More generally, this type of regression is far from ideal as a means of testing for welfare migration. For example, instead of relying on aggregate migration flows, one of the best welfare-migration studies for the US (Blank, 1988) uses individual relocation data for single mothers with children, a group likely to be welfare recipients. Blank relies on a multinomial logit framework to model the mother's location choice among 12 regions, whose average welfare benefit levels, unemployment rates, and other characteristics serve as explanatory variables. The results show a tendency toward relocation to high-benefit regions.

Another explanation for the small estimated interaction effect pertains to the uses of outdoor relief. While the New Poor Law authorities intended outdoor relief to be unavailable to able-bodied men and their families, this intent was often violated in practice, as explained earlier. Many localities, especially before 1835, persisted in relieving the un- and underemployed through cash assistance, both because they considered it a more cost-effective and humane alternative to indoor relief and because it allowed both rural and urban Poor Law Unions alike to manage cyclical labor demand.³⁶ Specifically, it allowed local employers, who in most Unions at this time dominated local welfare administration, to retain a pool of cheap labor nearby during slack periods (Boyer, 1993; Kiesling, 1996). In so doing, employers could then shift a fraction of their labor costs to taxpayers (Boyer, 1993). Higher outdoor benefits may then have prompted in-migration of the able-bodied, benefiting

³³ Implied net migration is calculated as the population count in the 1851 Census, less the population count in the 1861 Census, less the sum of annual deaths between these two censuses, plus the sum of annual births between these two censuses. The latter two figures are taken from the Registrar General's annual reports (this calculation yields the negative of the "error of closure"; see Arthi et al. (2017b) for more discussion). The net migration figure is then divided by the initial (1851) population in order to yield a rate.

³⁴ The other estimated coefficients in Table 6 are also of interest, and they show that in-migration was high in large, non-big-city districts with high shares of white-collar and nontradable employment, and low shares of other employment.

³⁵ An attempt to correct for endogeneity of **epp** was, however, not workable. When spatially lagged values of the covariates in Table 6 were used as instruments for **epp**, the Sargan statistic rejected their validity.

³⁶ Lees (1998) estimates that over the 1850–1870 period, about 10% of the population was in receipt of some relief in any given year, with this figure rising to 25% over a 3-year period. The majority of these paupers were granted outdoor relief (see Boyer (2002) for further discussion).

Table 6 Migration regressions

Variables	(1) <i>netmig_rate</i>	(2) <i>inmig_rate</i>
epp_tot	0.00154 (0.00197)	0.00117 (0.00105)
pop	0.00992** (0.00144)	0.00300** (0.000766)
bigcity	-0.107* (0.0528)	-0.0608* (0.0281)
cottondist	-0.00910 (0.0278)	-0.00419 (0.0148)
rateable_value	-0.0221** (0.00500)	-0.00384 (0.00266)
pop_shr_under_5	2.539** (0.742)	0.535 (0.394)
emp_shr_wc	2.054** (0.543)	0.875** (0.288)
emp_shr_nontrade	1.085** (0.308)	0.381* (0.164)
emp_shr_manuf	-0.190 (0.127)	-0.0379 (0.0676)
emp_shr_trans	-1.826** (0.669)	-0.507 (0.355)
emp_shr_other	-0.516** (0.179)	-0.305** (0.0949)
Constant	-0.506** (0.179)	-0.184 (0.0953)
Observations	536	536
R ²	0.580	0.573

Standard errors in parentheses ** $p < 0.01$, * $p < 0.05$

local businesses, while also encouraging in-migration of the destitute, who constituted a fiscal burden. This mixed migration picture, with its combination of positive and negative elements, may have weakened the predicted interaction in the choice of overall benefits.

While this argument seems plausible, it can be evaluated formally by allowing property owners to gain from the presence of poor residents who receive welfare. In the model, such an effect can be generated by assuming that the income y of property owners, instead of being fixed, is derived from profit in the production process where the poor are employed. An additional poor resident, by raising profit, thus confers a benefit on the owners while also entailing an extra tax burden. The model containing this amendment is analyzed in the appendix, and it reveals how the slope of the reaction function is affected. The slope increases, showing that compared to the case where their incomes are fixed, interaction is stronger, not weaker, when property owners gain profit from additional poor residents. Therefore, the

previous argument is not confirmed theoretically, although this conclusion comes from a highly stylized model that may fail to capture all elements of the Poor Law's environment.³⁷

7 Interaction in indoor and outdoor relief?

While strategic interaction seems to be present in the choice of total relief spending per pauper, did districts also interact in separately choosing the levels of its two components, indoor and outdoor relief? The answer appears to be negative, as follows. The data show outdoor spending per outdoor pauper, denoted **epp_out**, and after apparent errors in a handful of observations are eliminated via imputation, the mean values for this variable are near £ 4.5 in both periods, a low magnitude that indicates the relative cheapness of outdoor relief (recall that the **epp** means are around £ 7). Pauper counts show that around 87% of the total received this kind of relief. However, use of **epp_out** and its spatial lag in place of **epp** and **epp_lag** in the previous regressions leads to a statistically insignificant interaction effect. Next, the **epp** and **epp_out** variables along with counts of total and outdoor paupers can be used (after some further imputations) to compute **epp_in**, indoor spending per indoor pauper, whose means in the two periods lie in the £ 28–30 range, being much higher than the **epp_out** value. Reaction functions based on this variable again show insignificant interaction coefficients. Although this pattern is somewhat surprising, it suggests that in comparing their relief spending to that of their neighbors, districts focused on the overall package rather than its separate components. This view is made plausible by the fact that the split of total spending between indoor and outdoor relief varied greatly across districts, suggesting that total spending rather than its components offers a better picture of relief effort. The outdoor share in total spending ranges between approximately 0.09 and 0.91 in both periods, with means near 0.56 and standard deviations near 0.13 in both periods.

8 Conclusion

By using historical data to measure strategic interaction among local governments more than a century ago, this paper makes a notable contribution to an important area of research in public economics. The evidence, which shows the existence of a modest degree of interaction among British districts under the 19th-century Poor Law, helps to show that modern models of government behavior have external validity for periods in the past, when migration frictions may have been substantially greater. This finding provides an interesting counterpoint to present-day

³⁷ Clark and Page (2019) argue against Boyer's (1993) employer-benefit argument, in line with this theoretical disconfirmation.

concerns about the mobility of capital and taxpayers across and within countries in a globalized world and its effects on policy choices, showing that similar phenomena appear to have been at work in the distant past in a much more parochial setting.

In another departure from the literature, the empirical model allows the strength of interaction to grow following a welfare policy reform during the sample period, although the evidence does not indicate the presence of such an effect. However, the prediction of a post-reform drop in benefit levels is confirmed, although the drop is not large. While the study thus further validates modern expectations that local governments interact in their policy choices, it also advances the historical literature on poor relief in industrializing Britain. The relatively small quantitative effects that are estimated suggest that cross-jurisdictional differences in welfare benefits, while mattering, may have been just one of several factors affecting the location choices of migrants and the policy calculus of local officials.

The evidence of modest 19th-century welfare competition suggests that renewed study of this phenomenon in modern settings may be worthwhile. The 1996 welfare reform in the US replaced the AFDC program (Aid to Families with Dependent Children) by TANF (Temporary Assistance to Needy Families), with eligibility narrowed through various time limits on years of support. The study of Kaestner et al. (2003) suggests that these TANF restrictions reduced the extent of US welfare migration, perhaps indicating that welfare competition among the states (studied by Figlio et al. (1999) and Saavedra (2000) under AFDC) should have mostly subsided. Even so, a new US study for the TANF era could be useful along with studies for other countries, following Dahlberg and Edmark's (2008) work on Sweden.

Appendix

This appendix considers an extension where the income of property owners consists of profit from the production process in which the poor workers are employed. Profit in community 1 is equal to $\pi = f(n_1) - n_1 w_1 = f(n_1) - n_1 f'(n_1)$, and with profit divided equally among the owners, income is equal to $y = \pi/M$. The derivative of income with respect to E_1 equals

$$\frac{\partial y}{\partial E_1} = \frac{\partial y}{\partial n_1} \frac{\partial n_1}{\partial E_1} = \frac{-n_1 f''(n_1)}{M} \frac{\partial n_1}{\partial E_1} = \frac{\beta n_1}{M} \frac{a}{2\beta} = \frac{a n_1}{2\beta}, \quad (a1)$$

where the second to last equality uses (2) and $f'' = -\beta$. Substituting for n_1 using (2),

$$\frac{\partial y}{\partial E_1} = \frac{aN}{4M} + \frac{a^2}{4M\beta} E_1 - \frac{a^2}{4M\beta} E_2 \equiv G + KE_1 - KE_2. \quad (a2)$$

Rearranging (5), the previous first-order condition for choice of E_1 is $\Phi E_1 - \Omega E_2 - \Lambda = 0$. With income now dependent on E_1 , $\partial y/\partial E_1$ must be added to this expression, yielding

$$(\Phi + K)E_1 - (\Omega + K)E_2 - (\Lambda - G) = 0. \quad (a3)$$

Solving for the reaction function then yields

$$E_1 = \frac{\Lambda - G}{\Phi + K} + \frac{\Omega + K}{\Phi + K}E_2. \quad (a4)$$

When income y is a constant independent of E_1 , $K = G = 0$. Therefore, the effect on the reaction function's slope of moving to a positive K can be found by differentiating $(\Omega + K)/(\Phi + K)$ with respect to K . This derivative is proportional to $\Phi - \Omega$, which can easily be seen to be positive using (6) and (7). Therefore, when the income of property owners comes from profit rather than being fixed, the slope of the reaction function increases.

Acknowledgements I thank David Agrawal, Thushyanthan Baskaran, George Boyer, Jonathan Chapman, Matz Dahlberg, Kangoh Lee, Raphael Parchët, Luz Saavedra and several referees for helpful comments. Any shortcomings in the paper, however, are my responsibility. I am indebted to my UCI colleague Vellore Arthi and to her coauthors Brian Beach and W. Walker Hanlon for access to their data on the Poor Law and the demographics of mid-19th-century Britain (see Arthi et al., 2017a). I also thank Professor Arthi for extensive discussions and other input. In addition, I am indebted to Vanessa Nadalin for assistance in generating one of the weight matrices used below.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agrawal, D. R. (2015). The tax gradient: Spatial aspects of fiscal competition. *American Economic Journal: Economic Policy*, 7, 1–29.
- Agrawal, D.R., Breuille, M.-L., Le Gallo, J., (2021). Tax competition with intermunicipal cooperation. Unpublished paper.
- Allen, R. (2015). The high wage economy and the Industrial Revolution: A restatement. *Economic History Review*, 68(1), 1–22.
- Arthi, V., Beach, B., Hanlon, W.W., (2017a). Deportation risk, public health, and local labor markets: Evidence from the English Poor Law. Unpublished paper.
- Arthi, V., Beach, B., Hanlon, W.W., (2017b). Estimating the recession-mortality relationship when migration matters. NBER Working Paper No. 23507.
- Arthi, V., Beach, B., Hanlon, W.W., (2020). Recessions, mortality, and migration bias: Evidence from the Lancashire cotton famine. Mimeo.
- Baskaran, T. (2014). Identifying local tax mimicking with administrative borders and a policy reform. *Journal of Public Economics*, 118, 41–51.
- Besley, T., & Coate, S. (1992). Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs. *American Economic Review*, 82, 249–261.

- Blank, R. M. (1988). The effect of welfare and wage levels on the location decisions of female-headed households. *Journal of Urban Economics*, 24, 186–211.
- Boustan, L., Fishback, P., & Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the Great Depression. *Journal of Labor Economics*, 28, 719–746.
- Boyer, G. (1993). *An Economic History of the English Poor Law, 1750–1850*. Cambridge University Press.
- Boyer, G., (2002). English Poor Laws. EH.Net Encyclopedia, edited by Robert Whaples. May 7, 2002. URL <http://eh.net/encyclopedia/english-poor-laws/>.
- Boyer, G. (2019). *The winding road to the welfare state: Economic insecurity and social welfare policy in Britain*. Princeton University Press.
- Brueckner, J. K. (2000). Welfare reform and the race to the bottom: Theory and evidence. *Southern Economic Journal*, 66, 505–525.
- Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International Regional Science Review*, 26, 175–188.
- Chapman, J. (2018). Democratic reform and opposition to government expenditure: Evidence from nineteenth-century Britain. *Quarterly Journal of Political Science*, 13, 363–404.
- Chapman, J., (2020a). Extension of the franchise and government expenditure on public goods: Evidence from nineteenth-century England. Mimeo.
- Chapman, J., (2020b). Democracy, redistribution, and inequality: Evidence from the English Poor Law. Mimeo.
- Clark, G., & Page, M. (2019). Welfare reform, 1934: Did the New Poor Law in England produce significant economic gains? *Cliometrica*, 13, 221–244.
- Dahlberg, M., & Edmark, K. (2008). Is there a “race-to-the-bottom” in the setting of welfare benefit levels? Evidence from a policy intervention. *Journal of Public Economics*, 92, 1193–1209.
- Eugster, B., & Parchët, R. (2019). Culture and taxes. *Journal of Political Economy*, 127, 296–337.
- Feldman, D. (2003). Migrants, immigrants and welfare from the old Poor Law to the welfare state. *Transactions of the Royal Historical Society*, 13, 79–104.
- Figlio, D. N., Kolpin, V. W., & Reid, W. (1999). Do states play welfare games? *Journal of Urban Economics*, 46, 437–454.
- Fraser, D., (2009). *The Evolution of the British Welfare State* (4 ed.), Basingstoke: Palgrave/Macmillan
- Gibbons, S., & Overman, H. G. (2012). Mostly pointless spatial econometrics? *Journal of Regional Science*, 52, 172–191.
- Humphries, J., & Weisdorf, J. (2019). Unreal wages? Real income and economic growth in England, 1260–1850. *Economic Journal*, 129, 2876–2887.
- Horrell, S., Humphries, J., Weisdorf, J., (2020). Working for a living? Women and children’s labour inputs in England, 1260–1850. Mimeo .
- Kaestner, R., Kaushal, N., & Van Ryzin, G. (2003). Migration consequences of welfare reform. *Journal of Urban Economics*, 53, 357–376.
- Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17, 99–121.
- Kiesling, L. L. (1996). Institutional choice matters: The Poor Law and implicit labor contracts in Victorian Lancashire. *Explorations in Economic History*, 33, 65–85.
- Lees, L. (1998). *The Solidarities of Strangers: The English Poor Laws and the People, 1770–1948*. Cambridge University Press.
- Lindert, P. (1998). Poor relief before the welfare state: Britain versus the continent, 1780–1880. *European Review of Economic History*, 2, 101–140.
- Lloyd, A.J., (2007). Emigration, immigration and migration in nineteenth-century Britain. British Library Newspapers. Detroit: Gale.
- Lyytikäinen, T. (2012). Tax competition among local governments: Evidence from a property tax reform in Finland. *Journal of Public Economics*, 96, 584–95.
- Parchët, R. (2019). Are local tax rates strategic complements or strategic substitutes? *American Economic Journal: Economic Policy*, 11, 189–224.
- Patacchini, E., & Zenou, Y. (2012). Juvenile delinquency and conformism. *Journal of Law, Economics, and Organization*, 28, 1–31.
- Revelli, F. (2005). On spatial public finance empirics. *International Tax and Public Finance*, 12, 475–492.

- Rose, M., (1976). Settlement, removal and the New Poor Law. In: Fraser D. (eds) *The New Poor Law in the Nineteenth Century. Problems in Focus Series*. London: Palgrave.
- Saavedra, L. A. (2000). A model of welfare competition with empirical evidence from AFDC. *Journal of Urban Economics*, 47, 248–279.
- Southall, H., Gilbert, D., Gregory, I., (1998). Great Britain Historical Database: Labour Markets Database, Poor Law Statistics, 1859-1939. [computer file]. UK Data Archive [distributor] SN: 3713.
- Wheaton, W. C. (2000). Decentralized welfare provision: Is there a “race to the bottom”? *Journal of Urban Economics*, 48, 536–555.
- Wildasin, D. E. (1991). Income redistribution in a common labor market. *American Economic Review*, 81, 757–774.
- Wilson, J. D. (1999). Theories of tax competition. *National Tax Journal*, 52, 269–304.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.