The Role of Information in Explaining the Lack of Welfare-Induced Migration

Jeremy McCauley

Discussion Paper 20 / 729

22 October 2020

School of Economics
University of Bristol
Priory Road Complex
Bristol BS8 1TU
United Kingdom

The Role of Information in Explaining the Lack of Welfare-Induced Migration*

Jeremy McCauley†

This Version: 31st January 2020

Abstract: This paper examines the role of information as a driver of domestic welfare-induced migration decisions. I exploit a policy reform in England where the government began publicly releasing quality star ratings for each area's social services (social care). I study the effects of this "information shock" on the main service users, the elderly, and find a one star increase in publicly-released rating is associated with a 1.3% increase in the elderly population of that area. Based on empirical evidence, I estimate a search model with learning, where the elderly search for areas with better social services and gradually learn their true quality. Mimicking the information shock, the model reveals that individuals are more influenced by the rating of their own area than by ratings for other areas. Those induced to move by the information shock experience welfare increases valued at £600 per year.

Keywords: Local Government Differences, Regional Migration, Information, Search, Elderly
JEL codes: H73, R23, D83, J14

---

*I am grateful to Eric French for invaluable advice and guidance, and to Fabien Postel-Vinay for insightful suggestions and discussions. For helpful comments I thank Wei Cui, David Goll, Thomas Hoe, John Kennan, Rory McGee, Magne Mogstad, Cormac O'Dea, Ian Preston, Jeff Rowley, Peter Spittal, and Marcos Vera-Hernandez. Funding from ESRC Grant Inequality and the insurance value of transfers across the life cycle (ES/P001831/1) is gratefully acknowledged

†University of Bristol. Department of Economics. Email: jeremy.mccauley@bristol.ac.uk

1 Introduction

A large literature assesses whether geographic differences in the generosity of welfare programs within a country induces internal migration for those seeking more generous benefits, often called "welfare migration". Most studies find no effect or a very modest one.[1] Given that differences in welfare benefit generosity between areas can often be substantial this result is surprising.[2]

One explanation for why there does not appear to be much welfare migration despite the existence of differences in welfare benefits is that individuals are not aware of these differences. This paper examines the role of information as a driver of welfare-induced migration decisions, by studying how individuals respond to information releases regarding the quality of local public services and benefits. The literature to date studying information releases regarding the quality of local public services focuses on impacts to house prices.[3] To the extent that changes in house prices are driven by shifts in demand, these studies provide indirect evidence on the effect of information releases on migration. However, this is the first paper to *directly* consider what role information plays in migration decisions.

In the UK, personal social services (or social care) provision is decentralized to local authorities. The main users are the elderly, and benefits include services such as home help, home-delivered meals, nursing and residential homes, and day care facilities. Local authorities have a considerable degree of autonomy in the share of total revenues they assign to social services, in the quality of services they provide, and in any user-fees they charge. This creates large variations in the spending and quality of social services across areas. Around 0.5% of GDP was spent on social services for the elderly in 2001, but there was a wide range in spending across local authorities, with the highest spending around four times more (per capita) than the lowest.

I use a policy introduced into England in 2002 called the Social Services Performance Review (SSPR), where the national government gave a publicly-released star rating of each local authority's social services on a scale from zero to three stars based on a series of accounting and performance measures. I focus on the elderly, and in particular claimants of public (state) pension, who I refer to throughout this paper as pensioners. I treat this public release of the SSPR ratings as an "information shock" and analyze the number of

[1] Brueckner (2000) provides an excellent overview of this literature.

[2] Kennan and Walker (2010) look at the Aid to Families with Dependent Children program in the United States and find that even large differences in benefit levels provide surprisingly weak migration incentives for young welfare-eligible women. Similarly, Schwartz and Sommers (2014) find no significant migration effects on low-income people following recent expansions of Medicaid in certain states.

[3] For example, Figlio and Lucas (2004); Fiva and Kirkebøen (2011); Carrillo et al. (2013); Imberman and Lovenheim (2016); Haisken-DeNew et al. (2016); Hussain (2017) all consider at the effect of releasing school quality information on house prices.

pensioners in each area before and after the release, based on the rating that area received. The hypothesis is that if information does play a role then, after the information is made public, we would expect areas that receive higher star ratings to see increases in the size of the pensioner populations compared to the areas which perform poorly in the ratings. I find that a one star increase in publicly-released rating leads to a 0.01 percentage points increase in the percentage of pensioners living in that area relative to other areas. This corresponds to a 1.3% increase. For robustness I exploit the fact that the ratings were only conducted in England, whereas in Wales, which has the same local government structure, no ratings were produced.[4] Using predicted ratings, the Welsh local authorities are used as a control group for the English local authorities. With the control group included, the estimates remain very similar, which provides confidence in the results.

I use the results from the empirical investigation to motivate and estimate a migration model with learning. I draw inspiration from the literature on equilibrium search and unemployment, and in particular Moscarini (2005) and Papageorgiou (2014), but instead of workers searching for jobs, pensioners are searching for areas that provide better social services. The economy is fully populated with pensioners, who receive utility from only two sources: a public pension (which is not area specific), and social services (which is area specific). Each pensioner has an individual-specific quality "match" for the social services in each area, however this is unobservable to the pensioner. Social services are experience goods, which are a function of the match quality and observed with noise by pensioners. While the true match qualities are unknown to pensioners, they have beliefs about the quality of every area and update their beliefs based on the observed services received. Each period, pensioners assess their beliefs and decide if and where to move. An "information shock" is introduced into the model to replicate the introduction of the SSPR ratings by allowing, at the time of the shock, pensioners to receive information regarding the true mean qualities of every area. I estimate the model using indirect inference, and target moments characterizing migration, the distribution of pensioners across areas of different qualities, and regression estimates, mimicking the main empirical results in the paper.

Estimates suggest that there is a lot of noise in the learning process, and that it takes pensioners a long time to learn the true quality of their area. The information shock is met with an increase in net migration for the best areas and a decrease in net migration for the worst. The model reveals that when the star ratings were released, individuals placed more weight on the ratings given to their current area than on ratings for other areas. The ratings release led to increases in welfare, worth £600 per year for those influenced by the information

[4] Lockwood and Porcelli (2013) use Wales as a control group for England when assessing the impact of how councils responded to a similar rating system.

shock. I show that the form of the information shock is important and can lead to different welfare and migratory responses.

The two primary contributions of this paper can be summarized as follows. First, I exploit a natural experiment to present the first causal evidence that migration is affected by information releases. I document the extent of the response and the estimates are shown to be robust by finding a plausible control group. Second, using the empirical evidence as a motivation I develop and estimate a search model, where pensioners search for the areas with the best social services and gradually learn about their unobserved quality. The model is very tractable and can provide insights into how beliefs, the speed of learning, and information shocks can affect migration. The model can be applied to a wide range of applications where individuals have imperfect information about areas. This has consequences for the current literature on migration, which in general assumes individuals have perfect information about the quality of area-based amenities. When individuals instead have imperfect beliefs, this affects the estimation of other parameters, such as moving costs.

The findings in this paper are not only useful for other researchers but have important implications for policymakers regarding information releases, especially those relating to the quality of local services or benefits. Releasing information on the quality of local public services in becoming increasingly common in many countries. For example, in the UK there are school league tables, and star rating systems for local hospitals, care homes, and GP (Physician) practices. This type of information is also becoming easier for individuals to access on the internet. My results highlight that individuals do respond to these information releases, and that what is conveyed in the releases can affect the magnitude of the response and result in different outcomes.

The paper proceeds as follows. Section 2 surveys related literature. Section 3 discusses the context, data and descriptive statistics of social services in England, and the release of the SSPR in 2002. Section 4 presents and critically assesses the identification strategy. Section 5 presents the main empirical findings, discusses their economic significance, and reports results from a number of robustness checks. Section 6 introduces and estimates a search model with learning. The final section offers some concluding remarks.

2 Related Literature

2.1 Mobility Patterns of the Elderly

Dowding and Mergoupis (2003), while not focusing specifically on the elderly, find that people care about council provided services when making migration decisions.[5]

[5]Specifically, in a survey of three English areas that of people who moved a medium-distance in the

Theory on elderly mobility patterns (e.g. Graves and Knapp, 1988) suggests that the spatially invariant incomes of the retired should lead to migration toward areas where the wage and rent compensation for amenities occurs primarily in the labor market, rather than in the land market. Empirical research has found results consistent with these theoretical expectations. For example, Chen and Rosenthal (2008), using the 1970–2000 US Census, find that couples near retirement tend to move away from places with favorable business environments and towards places with highly valued consumer amenities. The size and quality of the public sector has also been shown to be an important determinant of elderly migration.[6] However, none of these papers account for whether individuals are actually fully informed of the spatial differences in the quality of the areas.

2.2 Welfare Migration

The question of whether or not differences in social services (which can be viewed as a welfare benefit) may induce elderly to move, is linked to the idea of welfare migration. Most of the welfare migration literature explores the issue of welfare competition and "race to the bottom". The theory speculates that in the presence of welfare recipients' mobility, decentralized welfare is set strategically by each authority as they consider the generosity levels of neighboring authorities before setting their own to avoid becoming a welfare magnet, leading to a race to the bottom in generosity of welfare. The majority of the existing empirical studies on welfare competition looks for strategic interaction among US states in setting welfare generosity. Brueckner (2000) provides a survey of these studies, most of which suggest that strategic interaction does occur. Looking at England, Revelli (2006), Moscone et al. (2007) and Fernandez and Forder (2015) all identify spatial interdependencies in social services expenditure levels between neighboring local authorities.

However, evidence on whether people are induced to move in order to obtain more generous welfare benefits is less clear.[7] Again, Brueckner (2000) provides a survey of the empirical evidence which is mixed, with some studies finding small effects, and most others finding an absence of welfare-induced migration.[8] One recent paper to look at this issue is Schwartz and

preceding seven years almost ten percent listed "*poor council services*" as a reason for leaving their previous location.

[6]Conway and Houtenville (1998; 2001) explore whether the elderly migrate to states with government policies that treat them favorably. Using state-level migration data from the 1990 Census, they estimate out migration and in migration equations that suggest that the public sector is an important determinant of elderly migration.

[7]As Brueckner points out, strategic interaction (and thus the race to the bottom) amongst states can occur even without the presence of welfare induced migration. All that is required is the *perception* by state governments that more generous benefits may attract welfare migrants.

[8]Meyer et al. (1998) also provides a good summary of the literature. Some studies do find evidence of welfare-induced migration. For example, Gelbach (2004), finds that among women likely to use welfare in the

Sommers (2014) who study low-income residents of states that forgo the Affordable Care Act's expansion of Medicaid, and who would be eligible if they moved to a state that did choose to expand coverage. They use a difference-in-differences analysis of migration in expansion and control states, and find no significant effects on migration. However, it is not clear whether or not the general public were aware about the welfare differences across areas or the size of any differences.

Kennan and Walker (2010) provide a systematic analysis of welfare-induced migration. They estimate a job search model based on a modified version of Kennan and Walker (2011), whereby young welfare-eligible women search across states for the optimal wage. The women know the wage in the current location, but to determine the wage at other locations it is necessary to move there. In each location, welfare acts as a fallback option, and the value of this is known by individuals. Performing counterfactual analysis on the model, it is found that that equalizing welfare benefits has a negligible effect on migration, regardless of whether the national benefit is set at either the lowest or the highest state benefit level.

2.3 Information Disclosures

This paper is related to a broader literature on how people respond to information shocks and disclosures. Crépon et al. (2018) study the role of information shocks, such as notifications of future training, in the context of a job search model, and find strong empirical evidence of workers responding to notifications of future training. In an experimental setting, Wiswall and Zafar (2015) provide information to college students regarding the population distribution of earnings of different college major choices. They find that college students are substantially misinformed about population earnings, and revise their self-earnings beliefs in response to being provided with information. The revisions are systematically related to the informativeness of the signal, with students exhibiting larger revisions when the information is more specific. There is also a growing literature (e.g. Conlon et al., 2018) which collects and uses subjective expectations data gathered from surveys to understand decision-making under uncertainty. Manski (2004) provides a survey of the early literature on this topic and a discussion on the importance of measuring expectations.

A body of work focuses on information and the take-up on benefits, however these are generally national benefits, not affected by location choice. This work has mostly concentrated on analyzing how providing information on eligibility affects take-up (e.g. Mastrobuoni 2011; Liebman and Luttmer 2015), or if providing personal assistance along with information increases take-up (Bettinger et al., 2012). Matikka and Paukkeri (2016) study the impact

US, those who move tend to move to higher-benefit states. Similarly, McKinnish (2007), finds estimates that are consistent with the presence of welfare migration effects.

of a reform in Finland where information letters regarding eligibility were sent to a portion of the eligible population before the implementation of a pension program. They find clear evidence that the information letters significantly increased take-up and prompted pensioners to apply sooner.

A growing literature considers how information on the quality of health providers, often in the form of "report cards", influence consumers' demand for health plans and health care providers (e.g. Beaulieu 2002; Scanlon et al. 2002; Wedig and Tai-Seale 2002; Dranove et al. 2003; Jin and Sorensen 2006; Chernew et al. 2008; Wang et al. 2011; Werner et al. 2012; Santos et al. 2016). Within this literature, Dafny and Dranove (2008) consider how much consumers learn about the quality of the providers from these public report cards versus what they would have learned in their absence, through "market-based" learning. They find that consumers learn from both public report cards and market-based sources, with the market-based sources having a larger relative importance.

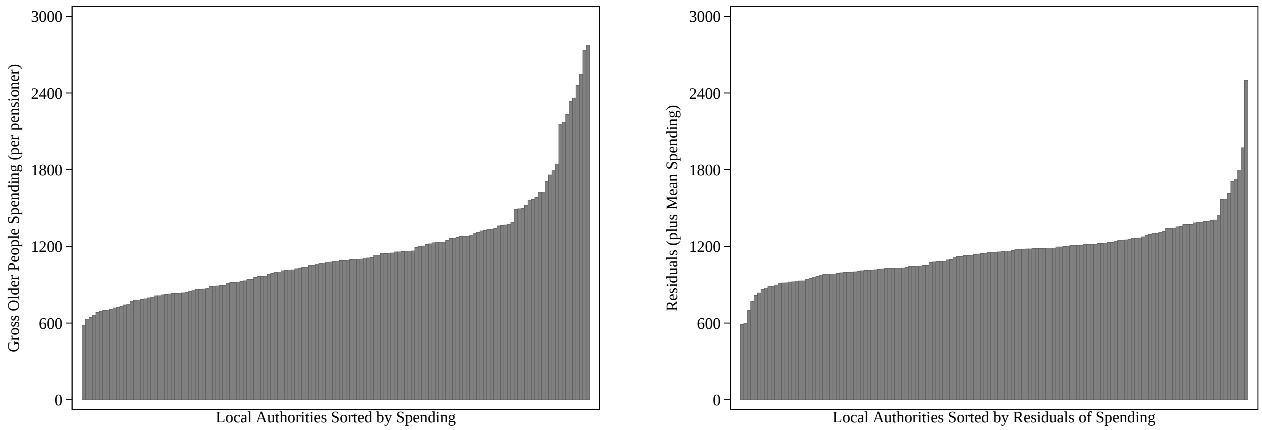
The majority of papers in the information disclosure literature consider the effect of information releases on house prices; looking at the effect of releasing information on school quality (Figlio and Lucas, 2004; Fiva and Kirkebøen, 2011; Carrillo et al., 2013; Imberman and Lovenheim, 2016; Haisken-DeNew et al., 2016) and even the effect of releasing information on earthquake vulnerability (Brookshire et al., 1985). To the extent that these changes in house prices are caused by shifting demand, most of these studies provide indirect evidence of the effect of information releases on migration. Some papers focus on how information on school quality can affect school-choice (Friesen et al., 2012; Hussain, 2017). However, this is the first paper to directly consider what role information plays in migration decisions.

3 Context, Data and Descriptive Statistics

3.1 Local Governments in England

3.1.1 Structure Local authorities in England are split into two main types of councils, unitary and two-tier. Each two-tier council is made up of a single county council with a number of lower-level district councils. Unitary and county authorities are responsible for the provision of social services, primary and secondary education, transport and waste disposal. There are 148 unitary and county authorities that cover the whole of England and do not overlap in social service provision.[9] In addition, there are 22 Welsh local authorities which perform the same functions.

[9] The Isles of Scilly and City of London are excluded due to their unique local government structures.



(a) Gross Spending on the Elderly (per pensioner) **(b)** Residuals of Gross Spending on the Elderly

Figure 1. Spending on Personal Social Services for the Elderly by Local Authorities in England, 2001

Notes: The figure displays the spending in 2001 of each of the 148 local authorities in England ordered along the x-axis by their spending. All amounts in 2015 £. Sub-figure 1b presents the residuals, $\epsilon_{i,2001}$, from the regression $y_{i,2001} = \gamma^{'} X_{i,2001} + \epsilon_{i,2001}$, where $y_{i,2001}$ is gross spending on the elderly (per pensioner) in local authority i. The area characteristics vector, $X_{i,2001}$, contains the percentage of population that are elderly; percentage of elderly that are disabled; percentage of elderly that are low-income; school quality; weather; political control; average wages; and average council tax.
Source: Local Government Comparative Statistics (LGCS) from the Chartered Institute of Public Finance and Accountancy (CIPFA).

3.1.2 Revenues and Expenditure In 2001, 37% of all local government revenues was from national government revenue support grants. Most of this grant revenue is not ring fenced, so it can be re-prioritized locally across services. The local authorities also use local taxes to complement the national government funds, with 22% of the revenue coming from local council tax. The largest area of expenditure for local authorities is education, which accounts for over 43% of their total expenditures in 2001. Personal social services is the second largest area of expenditure, accounting for almost 19% of total expenditure in 2001. Of the total expenditure on personal social services in 2001, almost 45% was spent on services for the elderly, with the rest split between services for the disabled and services for children and families. Across these user groups, supporting individuals into residential accommodation and nursing homes was the largest form of expenditure, accounting for 46% of total spending, compared to 39% for day care and home help provision.

Around £7 billion (2015 £) was spent by all local authorities on the elderly in 2001, but the range in spending across councils showed a wide variation. Figure 1 displays the ordered distribution of the local authority gross expenditure on personal social services in 2001. Panel (a) displays spending per capita of the elderly population, whereas panel (b) displays the residuals of spending per capita of the elderly population after controlling for a large number of area demographic characteristics. The figure highlights a heavy skew in the distribution, with the highest spending council spending almost four times more (per capita) than the lowest. The distribution remains unequal even after accounting for the demographic differences across councils. Fernandez and Forder (2015) show that the skew in the distribution persists even when spending is averaged across many years.[10]

3.2 Social Service Provision and the SSPR

In late 2001, then Secretary of State Alan Milburn announced the introduction of publicly-released star ratings for social services in England, called the Social Services Performance Review (SSPR). In April 2002, a letter was sent to the Directors of Social Services from the Chief Inspector of the which described how the ratings would be produced. According to the Social Services Inspectorate (SSI) the purpose of the SSPR ratings were to "*improve public information about the current performance of services, and the prospects for improvement*" (Social Services Inspectorate, 2002).

The SSPR ratings were to be similar to 'school league tables', which began in 1992 (Department for Education, 2018), and 'NHS hospital ratings' which began in 1995 (NHS Executive, 1995). Whereas those ratings provided information on the quality of education and hospitals in the area, the SSPR was to provide information on the quality of social services.

In 1995, the national government began measuring the performance of all local government services with the introduction of annual performance indicators (Audit Commission, 1995).[11] The indicators provided select comparative statistics for each local authority on the performance of all local services, including education and social services. The main purpose of this exercise was for the national government to ensure that local governments were getting value for money for their services. While some indicators were published, they were in general not easy for the public to access or interpret. The 2002 SSPR ratings used performance indicators related to social services, along with inspectors' judgments, to produce an overall star rating of the quality of social services in each area that was easy to interpret and widely publicized.

Councils were awarded either '3' stars ('excellent'), '2' stars ('good'), '1' star ('adequate') or '0' stars ('inadequate'). The best performing councils were given more freedom in the way they use national government provided grant funds. Councils with '0' stars were subject to more rigorous and frequent monitoring. The ratings that councils received were widely reported both in the local media and in the national media, with the BBC providing a dedicated web page listing the full results from every council ordered by star rating (BBC, 2002). Ratings were first published in May 2002, and were "refreshed" with additional

[10] Moscone et al. (2007) find a similar distribution for spending on services for mental health by local authorities in England.

[11] The performance indicators began being collected from the 1993/94 tax year and were renamed Best Value Performance Indicators (BVPIs) in 1997. BVPIs were discontinued in 2008.

information in November 2002 and then new ratings were released November of subsequent years. In this paper, the focus is on the initial ratings released in May 2002 as these ratings should have conveyed the most information to the public. Subsequent ratings could have been affected if more users moved to the area putting a strain on the service. It is also possible that after the initial ratings councils may have been able to “play” the system once they were aware what criteria were affecting their rating. Data on the SSPR ratings from 2002 come from the Department of Health’s SSI website.[12]

Figure 2a displays the 2002 SSPR ratings across local authorities in England. While a large set of performance evidence were used in constructing the star ratings, to try to ensure consistency a smaller set of Key Performance Indicators were also chosen. For these, a council could not be awarded the highest star rating if they failed to meet the “desired” level in even one of the Key Performance Indicators.[13] Spending levels were also factored into the star ratings, however higher spending did not necessarily lead to higher ratings, as some councils were penalized if they were deemed to be spending too much on certain services. This led, in some cases, to councils receiving fairly arbitrary star ratings. It is important to note that whether or not the star ratings convey information regarding the true underlying quality of social services in the area is not of prime importance to this study. What is more important is whether or not pensioners believe the information in the ratings and respond in their migration decisions.

A potential concern is that the SSPR star ratings may be associated with the quality of other local government services for which ratings exist and were published prior the SSPR ratings, such as ‘school league tables’. There is almost no correlation between the SSPR ratings received in 2002 and school quality variables such as the average GCSE points per pupil, and the percentage of schools put into special measures. Therefore, it is likely that the 2002 SSPR ratings presented new information to the public that was not previously known. Section 3.5 also provides some evidence that the star ratings that areas received do not appear to be associated with other area characteristics which may influence migration decisions.

[12]The Social Services Inspectorate was replaced by the Commission for Social Care Inspection in 2004 which was subsequently replaced by the Care Quality Commission in 2009. Their website can still be accessed through the UK Government Web Archive.

[13]There were a total of 11 Key Performance Indicators, which included statistics such as “Percentage of older people helped to live at home” and “Percentage of adults and older people receiving a statement of needs”.

3.3 Pensioners

Local authorities have some discretion in not only the level of service they provide but also the eligibility criteria.[14] Differing eligibility criteria may mean that, for example, some elderly individuals that are eligible for certain services in one local authority may not be eligible in another. In order to facilitate a consistent comparison across local authorities, instead of looking at the *service users* as defined by the local authorities, I focus on *potential users* by considering the population that are state pension claimants from the national government, of which eligibility is not dependent on receipt of local social services.[15] The state pension age during the time period of this study (1999-2006) was 60 for women and 65 for men. Take up of the state government benefits for the elderly is near universal, and should not be affected by a release of information on local social services in any way.[16] The focus of this study is how pensioners are distributed across the country, according to the quality of the social services in their area. The main dependent variable in this study is therefore the number of pensioners living in a local authority (as a percentage of all pensioners in the country).

3.4 Pensioners' Beliefs

Not much evidence exists on pensioners' beliefs regarding the quality of social services in their areas. Dowding (2005) attempted to use the 1997 Choice and Population Movement (CCPM) survey to address how changes in the quality of a given service affect citizens views about the performance of their local government. They asked participants "*At your current address, in general how satisfied or dissatisfied are you with your local council's provision of services?*"[17] These responses were then analyzed using ordered probit regressions on a wide range of output and performance measures for local authorities. Many of these performance measures are the

[14]The Fair Access to Care Services (FACS) guidelines were established by the UK Government in 2003 as a common framework for determining individuals' eligibility for social care services. However, local authorities continued to have some discretion in determining eligibility (Fernández and Snell, 2012).

[15]For this purpose of this study, I define *pensioners* as the population over state pension age who were claiming at least one of the key benefits: attendance allowance, disability living allowance, incapacity benefit, pension credit, state pension, and severe disablement allowance. The main national government benefit for the elderly in England is basic state pension, which is payable from the state retirement age. To be eligible for the full state pension individuals need to have paid national insurance contributions for 90% of working lifetime. Individuals with insufficient national insurance contributions are still able to receive a proportion of the state pension. The main national benefit for older disabled people is attendance allowance, which is for disabled individuals that require care. For an indepth overview of the UK benefits system during the time period this study focuses on see Emmerson and Leicester (2002) or Leicester and Shaw (2003).

[16]Emmerson and Leicester (2002) report that there were 10,963,000 basic state pension claimants in March 2001. However, the April 2001 Census reports that the UK resident population of men aged 65+ and women aged 60+ was only 10,810,878 (Office for National Statistics, 2001, Table P1). Part of this discrepancy can be explained by pensioners living abroad.

[17]Respondents could choose among: "very satisfied", "fairly satisfied", "neither satisfied nor dissatisfied", "fairly dissatisfied", "very dissatisfied", and "don't know".

same or similar to those later used to construct the SSPR ratings. Dowding (2005) could find no observable link between the stated responses and the performance measures. It is worth noting that the survey was conducted in 1997, which is five years before the SSPR ratings were publicly released.

3.5 Data and Descriptive Statistics

The analysis in this study is based on data from 148 English and 22 Welsh local authorities from several different sources between the years 1999 and 2006. Descriptive statistics (2001 means) and data sources of all the main variables used in the analysis are reported in Table 1. Full details of the data sources and can be found in the Data Appendix in section A.1.

The statistics in Table 1 are separated according to the area's 2002 SSPR rating, and show that councils separated by star ratings are similar across other mean control variables, such as the average weekly income. The '3' star councils have the highest house prices, however this is mainly driven by a couple councils located in London, and '2' star councils have the lowest average house prices. Even though house prices are highest in the '3' star councils, they have the lowest local taxes (council tax). Therefore the difference in the star ratings do not appear to have been driven by local taxes or other area characteristics.

Figure 2a displays the geographic variations in the SSPR star ratings received in 2002. There are a couple of things worth noting in the figure. First, there is no overall clear geographic pattern in the star ratings, especially concerning the lowest and highest ratings, where these councils are spread throughout the country. Secondly, comparing Figures 2a and 2b, which display the geographic variations in spending, there is no visible connection between the amount spent and the ratings received. Figures 2c and 2d display the geographic variations in how pensioners are distributed across the country and the percentage of each area's population that are pensioners. There does not appear to be any clear connection between the number of pensioners in an area and the star rating it received. Figure 2e displays the average house price, which are highest around London and the south of England. Figure 2f displays the net migration rate of the population aged over 60, which is the number of people who move into the area net of those who move out (as a percentage of that area's population). The figure shows that those aged over 60 tend to migrate away from London and the surrounding area and towards the North East and South West coasts of England. Importantly, the 2001 net migration for each area appears to have no connection to the 2002 SSPR ratings they received. To more formally test whether area characteristics can predict the 2002 SSPR ratings, in Table A2 in the appendix I regress a large number of area characteristic variables on the SSPR ratings in an ordered logit model. Out of 28 variables only one is statistically significant above the 10% level.

	Overall	By 2002 SSPR Rating				Source
		0	1	2	3	
Outcome variables						
Percentage of All						
Pensioners in a Given Council	0.676	0.679	0.613	0.796	0.575	D1
Total Percentage of Pensioners	100	6.790	49.653	39.004	4.600	D1
Number of Pensioners	60,392	60,680	54,746	71,141	51,363	D1
Population Aged Over 60	69,112	69,380	62,293	81,150	64,088	P1
Total Population	334,076	339,580	309,264	377,122	314,762	P1
Migration (Pop. Aged Over 60)						
In-Migration Rate (%)	1.555	1.311	1.570	1.588	1.501	P1
Out-Migration Rate (%)	1.848	1.807	1.971	1.605	2.141	P1
Net Migration Rate (%)	-0.293	-0.496	-0.401	-0.017	-0.640	P1
Average House Price (2015 £)	125,492	124,244	131,926	108,012	181,680	L1
Area Characteristics						
Average Weekly Income (2015 £)	954	929	978	917	983	A1
School Quality Measure	45.48	42.32	46.14	45.31	43.69	S1
Weather, Rainfall	66.59	62.18	66.75	67.36	65.70	W1
Weather, Average Temperature	10.18	10.24	10.23	10.02	10.48	W1
Political Control of LA						
Labour Dummy	0.426	0.400	0.432	0.449	0.250	E1
Conservative Dummy	0.230	0.200	0.185	0.245	0.625	E1
Lib. Dem. Dummy	0.047	0	0.074	0.020	0	E1
Other Party Dummy	0.007	0	0.012	0	0	E1
No Overall Control Dummy	0.291	0.400	0.296	0.286	0.125	E1
Average Council Tax	1,004	1,020	1,014	1,010	864	L2
Social Services Spending on Elderly	188	195	185	187	208	L2
Social Services Spending on Other	236	249	247	213	270	L2
Number of Local Authorities	148	10	81	49	8	

Table 1. 2001 Mean Local Authority Level Characteristics by 2002 SSPR Rating

Notes: The percentage of all pensioners in a given council is calculated as the number of pensioners in that council divided by the total number of pensioners in all councils. The in-migration rate is calculated as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in that area at the beginning of the year. The out-migration rate is calculated similarly for those who migrate out of an area. The net migration rate is the in-migration rate minus the out-migration rate. School quality is measured by the percentage of students who achieved at least 5 GCSE grades A*-C. Political control of the LA is the party that has a majority of the elected seats in the council. Social services spending is gross per capita (whole population).

Sources: A1: Annual Survey of Hours and Earnings

D1: From 5% samples of the DWP administrative data on the population over state pension age who were claiming at least one of the key benefits: Attendance Allowance, Disability Living Allowance, Incapacity Benefit, Pension Credit, State Pension, and Severe Disablement Allowance.

P1: Patient register data (1999-2008) from National Health Service Central Register (NHSCR).

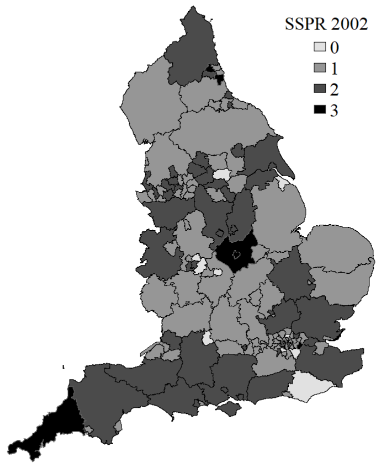
L1: Land Registry Price Paid data.

L2: Local Government Comparative Statistics (LGCS) from the Chartered Institute of Public Finance and Accountancy (CIPFA).

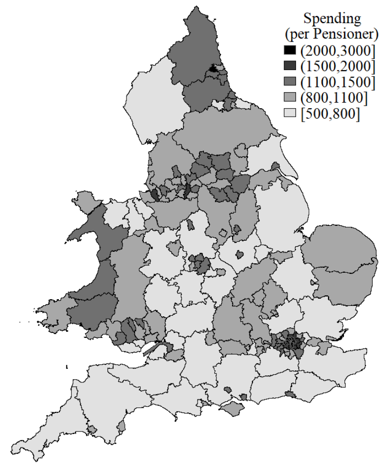
S1: GCSE results data from Best Value Performance Indicator (BVPI) data.

E1: Local council election results (1999-2006) are available on the BBC website.

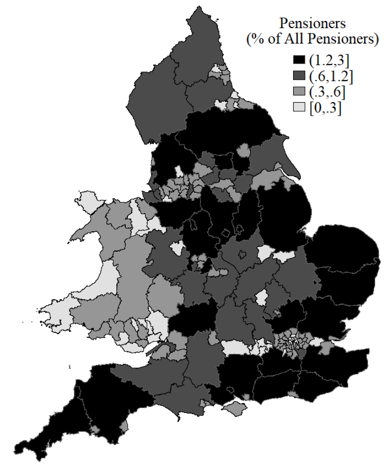
W1: Weather data for UK available from Centre for Environmental Data Analysis (CEDA).



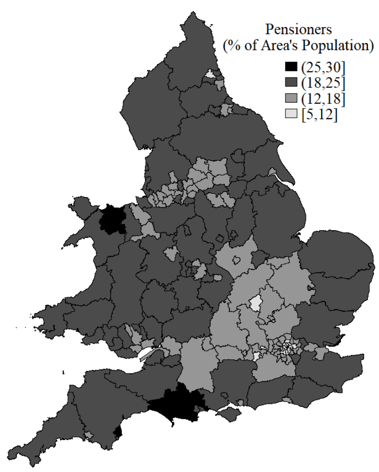
(a) 2002 SSPR Rating



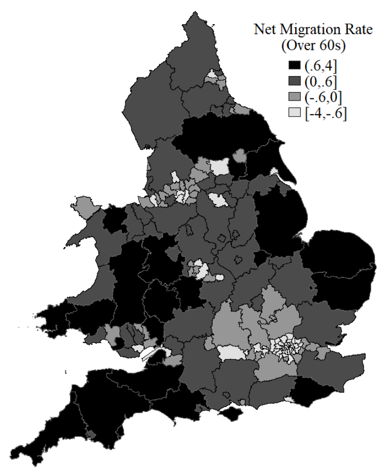
(b) PSS Spending (per Pensioner)



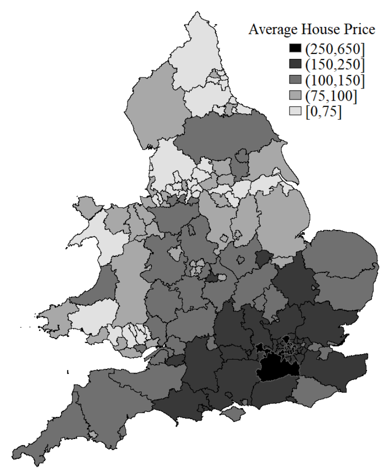
(c) Distribution of Pensioners Across Areas



(d) Pensioners (% Area's Population)



(e) Net Migration (%, Over 60s)



(f) Average House Price (000s)

Figure 2. Geographic Variations in 2001

Notes: This figure displays the 2001 geographic variations by local authorities (LAs) in England and Wales.

4 Empirical Strategy

The main estimates in this paper compare areas that received different star ratings, before and after this information was made public. Given that in 2002 the public were informed about the quality of social services in each area, those who would benefit from having better social services in their area, specifically pensioners, may react by moving to that area after mid-2002. I treat this "information shock" as a quasi-experiment. I define the 2002 public release of the SSPR as a binary treatment variable, switching on from 2002 onwards. The strategy compares the percentage of pensioners living in areas with different quality of social services (based on their 2002 SSPR star rating), before and after the public release of information. Treating the year 2002 as a cut-off point in the empirical analysis rests upon two assertions. Firstly, before 2002, there were common trends in knowledge and value of amenities across different quality councils. Second, the choice of using 2002 as a critical point in time is founded on the information in the publicly released SSPR ratings being widely spread.

The model can be written as follows:

$$Y_{it} = \beta_0 + \beta_1 P_t + \boldsymbol{\beta_D (R_i \times P_t)} + \gamma^{'} X_{it} + u_i + \tau_t + \epsilon_{it} \tag{1}$$

where Y_{it} is the outcome of interest for area i in year t, R_i is the 2002 SSPR rating from 0 to 3, P_t is a dummy which takes value 1 after SSPR was introduced (2002 onwards), X_{it} denotes area characteristics, u_i is area fixed effects and τ_t is year fixed effects. The coefficient of interest is β_D.

The set-up of this model is similar to differences-in-differences, however in this case there is no control group, as all areas received treatment (the information release). To address this, I use Welsh councils as a control group to analyze what would have been the path of the distribution of the pensioner population across English LAs, based on their underlying ratings, from 2002 onwards if the SSPR ratings would not have been publicly announced. Welsh councils are a suitable control group for the following reasons. First, Welsh and English LAs have the same structure and functions. Second, while the Welsh LAs have lower population densities than their English equivalents, the mean values of various control variables and quality measures are consistent across English and Welsh LAs, as can be seen in the appendix section A.3. Third, and crucially, while SSPR ratings were produced and publicly-released in England, no such ratings were produced for Welsh councils.

In the appendix section A.2, I predict the SSPR ratings that LAs in Wales would have received using accounting and performance measures, most of which were not public, in an ordered logit regression. A comparison between the predicted ratings (for the English LAs) and actual ratings can be seen in the appendix Figure A2. Reassuringly, only a few councils have predicted ratings different than their actual rating, and only ever by one star.

Using the Welsh LAs, separated by predicted ratings, as control groups for the English LAs of different ratings, the empirical approach is to estimate the impact of the introduction of SSPR in 2002 on the outcome Y_{it} as follows:

$$\begin{gathered} Y_{it} = \beta_0 + \beta_1 P_t + \\ + \beta_2 (P_t \times E_i) + \beta_3 (P_t \times \hat{R}_i) \\ + \boldsymbol{\beta_{DD} (P_t \times E_i \times \hat{R}_i)} \\ + \gamma^{'} X_{it} + u_i + \tau_t + \epsilon_{it} \end{gathered} \tag{2}$$

where E_i is a dummy which takes value 1 if the area is in England (the treatment group) and $\hat{R}_i$ is the predicted 2002 SSPR rating from 0 to 3. The coefficient of interest is β_{DD}.

4.1 Assessing the Identification Strategy

The fundamental identifying assumption is that time effects or trends are the same in the absence of the treatment (the SSPR ratings information release). In other words, the variable of interest should follow the same time path in each rating group in the absence of the treatment, conditional on other characteristics. This is not directly testable. However, here I provide some evidence that the common trends assumption cannot be rejected.

Visual inspections of the pre-reform years for all the main outcome variables as displayed in Figures A3 and A4 in the appendix show that, with only few exceptions, areas did not appear to have different trends before the reform according to what rating they received.

A more formal test that areas followed similar trends before the treatment can be conducted by running, for the pre-treatment period from 1999 to 2001, the following regressions:

$$Y_{it} = \tau_t + \theta_{d,t}(\tau_t \times R_i) + \gamma^{'} X_{it} + u_i + \tau_t + \epsilon_{it} \tag{3}$$

$$\begin{gathered} Y_{it} = \tau_t + \theta_{dd,t}(\tau_t \times R_i \times E_i) + \theta_{1,t}(R_i \times E_i) + \theta_{2,t}(\tau_t \times E_i) \\ + \theta_{3,t}(\tau_t \times R_i) + \gamma^{'} X_{it} + u_i + \tau_t + \epsilon_{it} \end{gathered} \tag{4}$$

where $\theta_{d,t}$ and $\theta_{dd,t}$ are the parameters of interest. Given that the social service ratings began in 2002, the null hypothesis that the variable of interest follows the same time path is simply $H_0 : \theta_{dd,99} = \theta_{dd,00} = \theta_{dd,01} = 0$. Table A4 in the appendix displays the p-values related to that null hypothesis for all of the main outcome variables considered in this paper for both of the above equations. In almost every case, the hypothesis that they did follow a common time path cannot be rejected.

Dependent Variable: Percentage of Pensioners in Area				
	Main Specification		Incl. Welsh Control	
	(1)	(2)	(3)	(4)
Information Effect of Star Rating				
Difference Term	0.0078**	0.0089***	0.0072**	0.0097**
	(0.0035)	(0.0029)	(0.0037)	(0.0041)
All controls		✓		✓
Number of LAs	148	148	170	170
Mean of Dep. Var.	0.676	0.676	0.588	0.588
Observations	1,184	1,184	1,360	1,360

Notes: The DD estimates in this table come from regression equation (1) and the DDD estimates from regression equation (2). The dependent variable is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country, as described in footnote 18. All regressions include LA fixed effects and year effects. Regressions in columns (1) and (2) are based on 8 years (1999-2006), for 148 English LAs. Columns (3) and (4) also include 22 Welsh LAs. Control variables are those listed under Area Characteristics in Table 1. Standard errors are clustered at the LA level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 2. Main Results: Distribution of Pensioners Across Areas

5 Main Empirical Results

This section presents the main results from the analysis. I start by considering how the release of the SSPR ratings in 2002 affected the distribution of pensioners across areas, in terms of the numbers of pensioners (as a percentage of all pensioners)[18], in areas that received different ratings.

I provide robustness checks. I then consider how the release affected different subgroups and population migration.

5.1 Distribution of Pensioners Across Areas

The main results use the distribution of pensioners across areas as the dependent variable, as described in footnote 18, in the equations (1) and (2). The estimates reported in Table 2 show that the release of star ratings in 2002 did have an impact on the distribution of pensioners across areas. The main estimates in column (2) suggest that a publicly-released one star rating increase is associated with a 0.009 percentage points (i.e. around 1.3%) increase in the

[18]The number of pensioners that live in that area as a percentage of all pensioners in the country is calculated for an area j as $100 \cdot \left(\frac{Number\ of\ pensioners\ in\ j}{\sum_s^J (Number\ of\ pensioners\ in\ s)} \right)$, where J is the total number of areas, which is 148 in England. DD regressions only include pensioners located in England, while the DDD specification includes all pensioners located in both England and Wales.

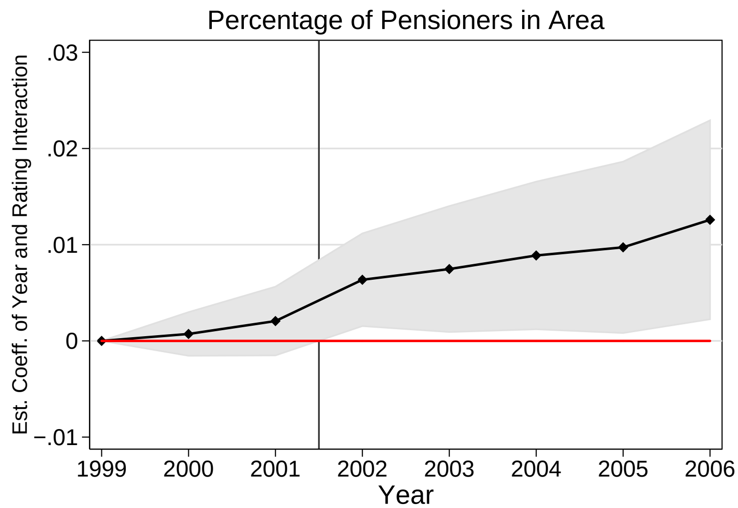


Figure 3. Visual Display of Impact of SSPR Star Ratings

Notes: This graph is a visual display of the estimates found in Table 2. This graph plots the β_{τ_t} coefficients for each year τ_t from the regression equation $Y_{it} = \beta_0 + \sum_i \boldsymbol{\beta_{\tau_t}(R_i \times \tau_t)} + \gamma' X_{it} + u_i + \tau_t + \epsilon_{it}$, where Y_{it} is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. Control variables are those listed under Area Characteristics in Table 1. The shaded gray area represents a 95% confidence interval around the estimates. Figure A6 displays a visual display of the estimates using predicted SSPR ratings, alongside the same graph for the Welsh control group.

percentage of all pensioners living in that area relative to others. The estimates in columns (3) and (4), where Wales has been used as a control group, support the main estimates with the point estimates remaining very similar in size, with a one star increase leading to a 0.008 percentage point (i.e. 1.4%) increase.

Figure 3 presents a visual display of the main estimates found in Table 2. The graph plots estimated regression coefficients of each year, of the sample years 1999-2006, interacted with the SSPR star rating. As can be seen in the figure, before the ratings were publicly released in 2002, the estimated coefficients are small and not statistically significant. The estimated coefficient jumps in 2002, corresponding with when the SSPR were publicly released.

5.2 Robustness

While the estimates that include the Welsh control group provide some confidence in the robustness of the main estimates, in this section I perform additional robustness checks, in particular with respect to some alternative methodological choices. I address inference and functional form dependence, and also perform a placebo test. The main robustness results can be seen in Table 3.

			Incl. Welsh Control	
	Estimate	Std. Error	Estimate	Std. Error
Base Specification	0.0089***	(0.0029)	0.0097**	(0.0041)
Including Trends				
LA-Specific Pre-reform Linear Trend	0.0089***	(0.0029)	0.0097**	(0.0041)
Excluding London LAs				
	0.0114***	(0.0042)	0.0130**	(0.0054)
Account for Possible				
Contamination from Welsh Movers	0.0089***	(0.0030)	–	–
Alternative variance estimators				
Bootstrapped Standard Errors	0.0089***	(0.0030)	0.0097**	(0.0042)
Collapsed for Two Periods	0.0102***	(0.0030)	0.00953**	(0.0041)
Alternative functional forms				
Log Transformation	0.0067*	(0.0036)	0.0147	(0.0112)
Non-Linear Estimation	0.0059***	(0.0010)	0.0057***	(0.0021)

Notes: The estimates in this table come from regression equation (1) and the estimates including the Welsh control group come from regression equation (2). The dependent variable is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. The Base Specification is column (2) in Table 2. All regressions include LA fixed effects, year effects and full controls unless otherwise stated. Control variables are those listed under Area Characteristics in Table 1. The number of observations for the specification excluding London LAs is 928, and 1,104 when including the Welsh control group. The Non-Linear column is estimated using a Pooled Bernoulli quasi-MLE. Standard errors are clustered at the LA level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 3. Further Robustness and Alternative Methodological Approaches

Additional controls The evidence presented in section 4.1 suggests that the introduction of the SSPR star ratings in 2002 is unrelated to most baseline area characteristics. Nevertheless, to examine whether the estimates are biased because of differential trends, I explicitly allow for differential LA-specific time trends. The LA-specific time trend is conducted by adding linear time trends extrapolated to the sample period based on pre-2002 data for each area. The estimates remain almost the same. Next, I check that the estimates are robust to excluding all LAs located in the capital city (London) and the results again remain very similar, but become larger in magnitude.

One potential issue with the main estimates is that even though Welsh LAs are not included in that specification, pensioners from Wales may still have responded to information shock regarding English LA star ratings and moved to English LAs, thus amplifying the estimates. To account for this I make use of migration data (for the over 60 population) and subtract any movers from Welsh LAs to English LAs in the counts of pensioners for each LA. The migration between Welsh LAs and English LAs is in most cases almost negligible and, as can been seen in Table 3, excluding these individuals does not change the main estimate.

Inference To ensure some outliers are not driving the statistical significance I bootstrap the standard errors. When this is done the standard errors remain essentially the same.

Bertand, Duflo, and Mullainathan (2004) point out that even with clustered standard errors, there can be downward bias in the standard errors, leading to false rejection of the null hypothesis of no treatment effect. To deal with this, I follow their recommended procedure of collapsing the time dimension to before and after the treatment, and re-estimating the model. This procedure produces very similar results to that of the baseline estimates in the paper.

Functional Form Dependence I check that the results are not dependent on the functional form by also estimating the model using a log transformation of dependent variable. I also estimate a non-linear version of the model using the methodology proposed by Papke and Wooldridge (2008) to tackle the possibility of non-linearity in case of fractional dependent variable. As the dependent variable lies between 0 and 1, I estimate a non-linear model as follows

$$Y_{it} = \Phi[\beta_d(R_i \times P_t) + \gamma^{'} X_{it} + u_i + \tau_t + \epsilon_{it}] + v_{it} \tag{5}$$

using a pooled Bernoulli quasi-MLE. Standard errors are clustered at the LA level, allowing for serial correlation in the v_{it}. Both cases support the baseline estimates as being robust.

I also estimate a version of the model where I allow for differential effects by each star ratings. The exact specification and the estimated results can be found in appendix section A.4. This specification also supports the idea that the release of SSPR ratings affected the composition of the population, and led to an increase in pensioners in the areas that got the highest ratings relative to others.

Placebo Figure 3 presents evidence that the information shock did occur in 2002, by showing a jump in the outcome occurring at 2002. To further assess to what extent the method is sensitive to picking up effects that are unrelated to the phenomenon in question, I shift the outcome measure relative to the information release, to act as a placebo test. This test is constrained by the limited number of observations in the pre-release period, however the public release of information should have no effect on the outcome in the years before the release. Shifting the outcome reduces the size of the effect and decreases the precision of the estimate, however the effect is still detected as future years still pick up the effect. In every

			Incl. Welsh Control	
	Estimate	Std. Error	Estimate	Std. Error
Pensioners	0.0089***	(0.0029)	0.0097**	(0.0041)
State Pension Only	0.0128***	(0.0044)	0.0124**	(0.0059)
Pension Credit (Low Income)	0.0283***	(0.0087)	0.0286***	(0.0108)
Disability Benefits	0.0070	(0.0047)	0.0089	(0.0060)
Benefit Claimants of Working Age	0.0039	(0.0030)	0.0042	(0.0043)
Income Related	-0.0032	(0.0045)	0.0009	(0.0051)
Disability Related	0.0033	(0.0047)	0.0008	(0.0069)
Unemployment Benefits (JSA)	-0.0358***	(0.0127)	-0.0378***	(0.0141)
House Prices (log)	0.0185	(0.0113)	0.0193	(0.0152)

Notes: The estimates in this table come from regression equation (1) and the estimates including the Welsh control group come from regression equation (2). Each row represents a separate regression. The dependent variable is the number of that group in the area as a percentage of the whole group. All regressions include local authority (LA) fixed effects, year effects and full controls unless otherwise stated. Control variables are those listed under *Area Characteristics* in Table 1. Standard errors are clustered at the LA level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 4. Distribution of Benefit Claimants Across Areas, By Different Types of Claimants

case, shifting the outcome variable backwards (or forwards) reduces the size of the estimate which indicates that 2002 is the correct treatment year.

Next, I test for a spurious relationship between the star ratings received and the evolution of the distribution of pensioners across areas by performing a random placebo test. Specifically, I randomly shuffle the star ratings of each council 2,499 times, keeping the overall true distribution of star ratings the same, and obtain the placebo coefficients from re-estimating the specifications from equations (1) and (2) as in Table 2. Figure A5 shows the distributions of the placebo coefficients and the original coefficients (marked by vertical dotted lines) along with kernel density plots. The distributions are centered at zero with the actual coefficients positioned to the very far right of the distribution. The areas to the right of the actual coefficients under the kernel distributions for the two specifications are 0.003 and 0.004, respectively.

5.3 Distribution of Other Potential Service Users Across Areas

There are reasons to expect that the release of the SSPR ratings may affect other subgroups of the population who may also be users of social services, such as the low income or the disabled. In this section I compare the effects on the pensioner population with other national government benefit claimants.

The top part of Table 4 subdivides pensioners according to what government benefits they are claiming. As can be seen the point estimate for low income claimants is over twice as large as those only claiming state pension (who are not low income). This is as would be expected, as the main beneficiaries of local social services would be low-income elderly. Perhaps surprisingly the estimates for pensioners claiming disability benefits is much smaller and not statistically significant. This is a group, along with low income pensioners, that would be gain a lot from local social services. This may point to the fact that those on disability related benefits are less mobile than others.

Table 4 shows that the release of the SSPR ratings had a stronger effect on pensioners than on working age benefit recipients. The point estimate for the working age benefit recipients is smaller at 0.004 percentage points and is not statistically significant. When the working age benefit recipients are separated into their main area of benefits (income related or disability related) it can be seen that most of the overall point estimate comes from the disabled, and the point estimate for those on income related benefits is actually negative. However, neither are statistically significant. Again, this may be caused by disabled claimants being mobility constrained.

As expected, and as an additional robustness check, the SSPR ratings release appears to have no positive effect on the number of unemployment benefit (i.e. jobseeker's allowance) claimants in the area, with the point estimate being negative. This is reassuring as there is little reason to think that those claiming unemployment benefits would be users of local social services, and instead may be attracted to councils that spend more on other services such as education or local business initiatives.

5.4 Migration

One common criticism of studies looking at welfare migration but focusing on the stock or composition of the population instead of migrants in or out of an area, is that of endogenous participation. Instead of picking up a welfare migration effect any estimates may instead pick up an effect of more individuals in an area starting to claim benefits (and not previous claimants moving into the area). This is particularly likely when a certain area introduces more generous benefits. This should not be a problem in this study, as the main outcome that is the percentage of the population claiming *national* state pension benefits, which is not dependent on the area they live in and whether or not they make use of local services. This segment of the population is the one that is most likely to benefit from local government social services, however the pensioners may or may not be eligible for these services depending on the area. Nevertheless, to ensure endogenous participation is not biasing the estimates, I check if the introduction of SSPR in 2002 had an impact on the actual migration into the area.

Indirect evidence on the migration response can be seen in the house price response. Given this information release should have made areas with higher star ratings more desirable we might see this increased demand for an area reflected in a change in house prices. The estimates in the bottom panel of Table 4 shows that a one increase in publicly-released star rating was related to a 1.9 % percent increase in house prices, however the estimates are not statistically significant.

The next set subsection considers estimates from migration data for the population aged 60 and older. These movers cannot be subdivided into whether or not they were state pension claimants, but the estimates will still shed light on the overall effect.

5.4.1 In-Migration, Out-Migration and Net Migration I define the in-migration rate as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in the LA at the beginning of the year. Similarly, the out-migration rate is the number of individuals that migrated out of an area by the end of a given year, divided by the size of the population in the LA at the beginning of the year. The net migration rate is simply the in-migration rate minus the out-migration rate. The estimates that follow only focus on the migration rates of the over 60 population.

Panel (a) of Table 4 shows that the release of the SSPR ratings appears not to have had much, if any, of an effect on the in-migration rate of the area, with the point estimates being negative and imprecise. However, there appears to have been a sizable effect on the out-migration rate of the area with a -0.034 percentage point decrease in the over 60 population out-migration rate for a one star increase in SSPR rating. This suggests that people are less likely to leave an area once they find out that the social services are good. Most of this affect appears to be coming from those aged 60-74, which may indicate that either those over 75 are less mobile or are less likely to respond to an information release.

The estimates shows that the net migration rate of an area as rose by 0.026 percentage points for a public release of a one increase in star rating, however this estimates are not statistically significant. This implies that the SSPR ratings release resulted in a larger increase

Migration	Panel (a): Overall Migration Rates by LA (%)					Panel (b): Migration Flow Rates Between LAs (%)				
				Incl. Welsh Control					Incl. Welsh Control	
	Mean	Estimate	SE	Estimate	SE	Mean	Estimate	SE	Estimate	SE
In-Migration Rate (Aged 60+)	1.407	-0.0074	(0.0138)	-0.0431	(0.0519)	0.0085	-0.0001	(0.0001)	0.0000	(0.0001)
Aged 60-74	1.362	-0.0156	(0.0167)	-0.0555	(0.0576)	0.0083	-0.0001	(0.0001)	0.0001	(0.0001)
Aged 75+	1.430	0.0047	(0.0179)	-0.0259	(0.0532)	0.0091	-0.0002*	(0.0001)	-0.0000	(0.0002)
Out-Migration Rate (Aged 60+)	1.671	-0.0335*	(0.0173)	-0.0110	(0.0324)	0.0102	-0.0003***	(0.0007)	-0.0004***	(0.0001)
Aged 60-74	1.615	-0.0458**	(0.0210)	-0.0324	(0.0393)	0.0095	-0.0004***	(0.0001)	-0.0004***	(0.0001)
Aged 75+	1.712	-0.0167	(0.0180)	0.0259	(0.0367)	0.0107	-0.0002**	(0.0001)	-0.0003*	(0.0001)
Net Migration Rate (Aged 60+)	-0.263	0.0261	(0.0242)	-0.0321	(0.0487)	-0.0016	0.0002*	(0.0001)	0.0004**	(0.0001)
Aged 60-74	-0.253	0.0303	(0.0284)	-0.0231	(0.0579)	-0.0016	0.0003**	(0.0001)	0.0005**	(0.0002)
Aged 75+	-0.282	0.0214	(0.0268)	-0.0518	(0.0460)	-0.0016	0.0001	(0.0001)	0.0003	(0.0003)
Fixed Effects		Year, LA		Year, LA			Year, LA-pair		Year, LA-pair	
Number of LAs		148		170			148		170	
Number of LA-Pairs		–		–			10,878		14,365	
Observations		1,184		1,360			174,048		229,840	

Notes: In panel (a), the estimates in this table come from regression equation (1) and including the Welsh control group estimates come from regression equation (2). In panel (b) the estimates come from a modified version of equation (1), which is given in equation (6). The estimates in panel (b) that include the Welsh control group come from the same modification for regression equation (2). The in-migration rate is calculated as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in that area at the beginning of the year. The out-migration rate is calculated similarly for those who migrate out of an area. The net migration rate is the in-migration rate minus the out-migration rate. In panel (a) the dependent variables are the overall migration rates for each LA, whereas in panel (b) the dependent variables are the the migration rates between LAs. All regressions include include local authority (LA) or LA-pair fixed effects, year effects and full controls unless otherwise stated. Control variables are those listed under *Area Characteristics* in Table 1. Regressions are based on 8 years (1999-2006). The main estimates include 148 English LAs, and the columns including Welsh control group adds 22 Welsh LAs. Standard errors in panel (a) are clustered at the LA level, and in panel (b) are clustered at LA-pair level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 5. Effects on Migration

in the size of the elderly population in areas that performed well than what was found in the main estimates in Table 2. However, the estimates are not directly comparable as Table 2 focuses on state pension claimants, whereas Table 4 is focusing on the whole population aged over 60.

5.4.2 Migration Between Areas While the results in Panel (a) of Table 4 appear to show that the outflow of the over 60s was affected by the information release, some of the results are imprecise and there does not appear to have been any effect on in-migration. The response may be weak due to some frictions between areas, such the distance between them. To investigate this possibility, I examine the flows between LA-pairs. In particular, equation (1) is adjusted by including an LA-pair fixed effects so that the equation becomes

$$Y_{ijt} = \beta_0 + \beta_1 P_t + \boldsymbol{\widetilde{\beta}_D(\widetilde{R_{ij}} \times P_t)} + \gamma' \widetilde{X_{ijt}} + f_{ij} + \tau_t + \epsilon_{it} \tag{6}$$

where Y_{ijt} is the migration rate between area i and j (divided by the population size of area i), f_{ij} is an LA-pair fixed effect (with $f_{ij} = f_{ji}$), $\widetilde{R_{ij}}$ is the difference in star ratings between area i and j, and $\widetilde{X_{ijt}}$ is the difference of the control variables between area i and j. The coefficient of interest in this specification is is $\widetilde{\beta}_D$. The interpretation of this estimate is how the migration rate between area i and j responds to the information being released regarding a one increase in the *difference* of star ratings across areas. Similar to before, equation 2 can be amended using Wales as a control group to calculate the estimate, $\widetilde{\beta}_{DD}$.

The results of this specification can be seen in Panel (b) of Table 5, and they line up with the estimates in Panel (a). In particular, a public release of a relative difference of one star in the SSPR ratings is associated with a reduction in out-migration of 0.0003 percentage points, which corresponds to a 3 % decrease in the outflow of pensioners from the area that received the higher rating to the area with the lower rating. Similarly, the release of SSPR is associated an increase in net migration. Again, it appears to be those aged 60-74 that are driving the estimates, with most of the point estimates for those aged over 75 being smaller and less precise.

6 Model with Imperfect Beliefs and Learning

So far, I have presented evidence that information does play a role in explaining the lack of welfare induced migration, by showing that the release of the SSPR ratings in 2002 affected pensioner migration. However, with the empirical evidence alone it is impossible to quantify the extent of the role that information plays, or to consider the impact of other factors such as moving costs. In particular, it is difficult to know how weight pensioners gave to the information shock. For example, it may have only taken very small changes in beliefs to generate the response in the empirical example, or alternatively it may have taken a large changes in beliefs. It is also important to establish whether or not the information release was good for pensioners in terms of utility. In this section, I present a model with imperfect beliefs and learning that will help answer some of these questions.

The model I present draws from the literature on equilibrium search unemployment, but instead of searching for jobs, pensioners are searching for the area that provides the best social services. The model, while on a different application, draws in particular from Moscarini (2005) and Papageorgiou (2014) who nest a job matching model into an equilibrium search unemployment framework. The model also draws from Pessino (1991), who analyzes a sequential migration model with imperfect information. Also highly related, while not in a search framework, is Chernew et al. (2008), who study the consumer choice of healthcare plans after a release of "report cards" in a Bayesian learning framework.

In my model, pensioners search for the areas which provide the best social services. Social services is an experience good which is measured with noise. Pensioners have beliefs over the unknown quality of social services in every area and gradually learn the unknown quality of their current area's social services using Bayesian updating based on the observed services received. Based on their beliefs about the quality of the services in the area they are in (their "match"), and their beliefs over the services in other areas, each pensioner decides when and where to move.[19] The set-up is most similar to Papageorgiou (2014), who builds a job search model with three occupations and workers who learn their productivities in each job over time. In my model, beliefs are not reset upon moving to a new match, but are kept (and updated) until an individual dies. This is similar to Papageorgiou (2014) and Eeckhout and Weng (2010), but different than most of the rest of the literature (e.g. Jovanovic, 1979; Moscarini, 2005; Decreuse and Tarasonis, 2016; Li and Weng, 2017). Therefore, in the model, if a pensioner leaves an area due to the belief that the social services quality match is poor, they will keep that same belief about that area unless they return and update their beliefs or they receive outside information (i.e. an information shock).

6.1 Population of Social Service Users

The economy is entirely populated by risk-neutral pensioners, who search for the areas which provide the best social services. Pensioners receive utility from two sources: public pension payments from central government (which are the same for all pensioners regardless of area) and local social services (which have individual-specific match qualities) provided by an

[19] Alternatively, this problem could have been set up as a reputation model (see Jolivet et al., 2016), with the councils being the seller and the pensioners being buyers.

area's local council. The individual-specific true match quality for each council is unknown. While the true qualities are unknown, pensioners have beliefs about the quality of *every* council (not just the one they are currently in). Pensioners can choose to move for a fixed council-specific moving cost. There is a probability of death, γ. Each pensioner who dies is replaced immediately by a new pensioner, keeping the population constant. The model is in discrete time and the timing of events is as follow. Consider a pensioner beginning the period in a council j:

1. With probability δ, they face an exogenous "family shock" and move to a (random) new council .
   (a) If they do not have a family shock, they will assess their beliefs regarding social services and decide if they want to move. If they choose to move, they will pick the council that they believe will offer them the highest quality. They will then move to that council, paying the council-specific fixed moving cost.
2. They receive a state pension payment from the central government, which is the same regardless of which council they are in.
3. They experience their local social services, which is specific to the pensioner and the council they are in. They then update their belief about the quality of their current council based on that experience. The beliefs about the councils they are not in remain unchanged.
4. With probability γ, they die.

If the pensioner does not die, the above steps are repeated.

6.2 Local Councils and Social Service Quality

In order to match with my empirical investigation earlier in the paper, in the model there are four different mean area council qualities: $\mu_0 \leqslant \mu_1 \leqslant \mu_2 \leqslant \mu_3$ which correspond to councils with '0', '1', '2' and '3' SSPR stars respectively. Assume, for now, that there are only four local councils for each pensioner to choose from, one of each quality type. Then for local council $j = 1, .., 4$, at birth into the model, pensioner i draws their unobserved individual-specific true qualities from the distribution:

$$q_{i,j} \sim N\left(\mu_j, \tau^2\right) \tag{7}$$

where τ determines the dispersion of individual qualities around the mean for each council. The individual-specific true qualities are time invariant. These match qualities, $q_{i,j}$, incorporate all of the individuals area-specific amenity preferences, however these cannot be separately identified from the preferences for social services.

The total per period utility received by pensioner i who is currently in council j is:

$$y_{i,j,t} = y^c + y^s_{i,j,t} \tag{8}$$

where $y^s_{i,j,t}$ is the observed local social services from council j at time t, and y^c is state pension payments from central government. The observed social services is an experience good, which is a function of the individual-specific true match quality but also has noise. Specifically, each period t the pensioner's observed social services is drawn from the distribution:

$$y^s_{i,j,t} \sim N\left(q_{i,j}, \sigma^2\right) \tag{9}$$

where σ determines the dispersion of the noise in the observed social services, which affects the variation in their social services experiences and will influence the speed at which pensioners learn.

6.3 Beliefs

Consider pensioner i, living in area j, at time period 1. They do not know the true value of their match quality, $q_{i,j}$, but they do know the true variance τ. Therefore, they assume the prior distribution of $q_{i,j}$ to be normal with mean $p_{0,i,j}$ and variance τ^2. The mean of their prior distribution is drawn for each area at birth into the model as follows:

$$p_{i,j,0} \sim N\left(\mu_j, \kappa^2\right) \tag{10}$$

where κ determines the dispersion of initial beliefs around the mean true qualities of the councils.

After observing their received social services each period, $y^s_{i,j,t}$, pensioners update their beliefs about the quality of that area using Bayes' rule. For example, assume that there is no discounting and that individual i does not move away from j. Then i's potential observation of social services at location j in time period 1 is:

$$y^s_{i,j,1} = q_{i,j} + \sigma\epsilon_{i,j,1}, \;\; \epsilon_{i,j,1} \sim N(0,1)$$

where the last term in the equation concerns the noise of the process. The conditional probability distribution of the process state given the available information, I, is $P(q_{i,j}/I_1)$, where $I_1 = 0$. That is, initially the individual has no additional information. Information in the second period, $I_2 = (y^s_{i,1})$, is based on their observations in the first period. The conditional distribution when the available information is I_1 is the normal distribution with mean $p_{0,i,j}$ and variance τ^2. Now, in the second period, the posterior distribution is also a normal distribution but with mean $\left[y^s_{i,j,1}\tau^2 + \sigma^2 p_{0,i,j}\right] / \left[\tau^2_{i,j} + \sigma^2\right]$ and variance $\tau^2_{i,j}\sigma^2 / [\tau^2 + \sigma^2]$.

This generalizes so that for any council j that pensioner i is currently living in, the mean and variance of the posterior distribution are updated across time periods according to following equations:

$$p_{i,j,t_j+1} = \frac{y^s_{i,j,t_j}(\hat{\tau}_{i,j,t_j})^2 + \sigma^2 p_{i,j,t_j}}{(\hat{\tau}_{i,j,t_j})^2 + \sigma^2} \tag{11}$$

$$\hat{\tau}_{i,j,t_j+1} = \hat{\tau}_{i,j,t_j}\sqrt{\frac{\sigma^2}{(\hat{\tau}_{i,j,t_j})^2 + \sigma^2}} \tag{12}$$

where t_j is the total length of time that a pensioner has lived in their current council j.[20] Initially, at $t_j = 0$, for each pensioner and area, the variance $\hat{\tau}_{i,j,0} = \tau$, but this value will get smaller the longer they live in a particular area and the more social services experiences they have. Over time, when updating their beliefs, pensioners will place less weight on new shocks to their observed social services versus their current beliefs. Note that each pensioner only updates the beliefs of the council they are living in, but has a memory of their current beliefs for every council. Therefore, $\hat{\tau}_{i,j,t_j}$ determines the accuracy of pensioner i's current belief in council j.

6.4 Value Functions

Let $\tilde{p}$ denote the vector of beliefs for each pensioner, and let $C(\tilde{p})$ denote their value function (excluding any moving costs) when they have the option of moving to any area, for ease of notation I suppress individual and time subscripts. Then,

$$\begin{aligned} C(\tilde{p}) = \quad & max\{y_1(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_1(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,, \\ & y_2(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_2(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,, \\ & y_3(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_3(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,, \\ & y_4(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_4(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,\} \end{aligned} \tag{13}$$

where $\bar{V}(\tilde{p})$ denotes the weighted average value function a pensioner would expect if they receive an exogenous move to a random council. If k_j denotes the fixed cost of moving to j from any council, then, for example, the value function of a pensioner currently living in a council with '0' stars is

[20] For each pensioner, as there are 4 councils, $t = t_1 + t_2 + t_3 + t_4$. If the pensioner has only ever lived in council j, then $t_j = t$.

$$
\begin{aligned}
V_0(\tilde{p}) = \quad & \mathbb{I}\left(argmaxC(\tilde{p}) = 0\right) C(\tilde{p}) \\
& +(1 - \mathbb{I}\left(argmaxC(\tilde{p}) = 0\right))[C(\tilde{p}) - k_{argmaxC(\tilde{p})} \\
& +\beta(1-\delta)(1-\gamma)\mathbb{E}V_0(\tilde{p}) + \beta\delta\mathbb{E}\bar{V}(\tilde{p})]
\end{aligned} \tag{14}
$$

The solution to this problem will reveal, for each individual based on their beliefs, which is the preferred council, and whether the they should move. However, especially when the number of areas increase, this problem does not have a straightforward closed-form solution. Therefore the solution needs to be derived numerically in a process which I describe in the next section.

6.5 Estimating the Model

I estimate the model by matching features of the model to the data from my earlier empirical investigation. To start, I simulate individual histories for a large number of pensioners. Time is discrete, where each period is a year. Whereas in the derivations above I assume there are only four councils, in the estimation there are 148 councils to mimic the number of councils in England. Each council has the mean council quality that correspond to the 2002 SSPR rating that they received.

While there are 148 councils in the model, to make estimation feasible, each pensioner has a set of only five areas from which they have beliefs and can choose to move.[21] The set of five areas that each pensioner has is determined by the initial random area in which they reside. After drawing their initial area, the four other areas are randomly drawn from a set of the ten most moved to areas from that initial area from the over 60s in the migration data. While this may appear restrictive, in the data 65% of all moves from a LA will be to at most ten other LAs, with the remaining 35% of moves from that LA spread over the remaining 137 LAs.[22]

All pensioners start in a random council with draws of their unobserved individual-specific true qualities and their initial beliefs of the qualities for every council in their set of five. The timing of events for each pensioner is as described in section 6.1. Each pensioner can move at most once during a period. If a pensioner dies, they are immediately replaced at the start

[21]The reason that the model is estimated with only five councils for each pensioner, and not more, is due to the fact that pensioners have beliefs about the quality of *every* council. Therefore, each pensioner has the same number of state variables as there are councils. Increasing the number of councils has a very large computational cost, which is discussed more in the appendix section A.5. Kennan and Walker (2011) drastically reduce the size of the state space in their job search model by reducing the information set to only include wages seen in locations recently visited by the individual. In my application, however, this technique is not appropriate as the information shock is going to convey information about every area.

[22]In the data, from 1999-2001, when moves from each LA are sorted according to the most migrated to areas to the least, on average 23% of all moves are to the single most popular area, 48% of all moves are to the four most migrated areas, and 65% of all moves are to the ten most migrated areas.

of the next period by a new pensioner in a random council. I simulate this for a long time period to ensure a steady state is reached, and then introduce an "information shock" to mimic the release of the SSPR ratings as in the empirical investigation.

I introduce an "information shock" as follows. When the shock hits in a certain period, pensioners gain an additional social services experience for every council (not just the one that they are currently in). The additional experience does not update beliefs in the ordinary way according to equation (11). Instead, the experience provides pensioners with information on the mean quality of each council. However, the amount of weight that pensioners place on the new information can vary according to whether or not it is regarding the council where they currently reside. Specifically, at the time of the shock for a pensioner residing in council j, beliefs are updated as:

$$p_{i,j,t} = p_{i,j,t-1} + \Delta_p \left(\mu_j - p_{i,j,t-1}\right) \tag{15}$$

and for all other councils $k \neq j$,

$$p_{i,k,t} = p_{i,k,t-1} + \Delta_{\bar{p}} \left(\mu_j - p_{i,k,t-1}\right) \tag{16}$$

This represents the fact that pensioners may receive more information regarding their current area than for other areas. For example, the majority of local newspaper stories concerning the SSPR would report on the local area's rating, but not mention the ratings of other areas in the country. It is also important to note that this shock does not necessarily bring beliefs closer to a pensioner's true values. This is due to the shock bringing beliefs closer to the true mean quality for each council, μ_j, not the true individual-specific values, $q_{i,j}$. This reflects the fact that the release of the SSPR ratings only gave councils a broad star rating, instead of giving pensioners personalized information. Pensioners that are born into the model after the shock has taken place have this experience immediately after birth. Sections 6.5.4 and 6.6.1 give a greater discussion on the form that the information shock takes, and consider the effects of an alternative shock which instead provides personalized information.

I estimate the model parameters using indirect inference (Smith, 1993; Gourieroux et al., 1993). Suppose that the parameters of the auxiliary model satisfy:

$$\bar{m} = m(X_N) = \arg\min_m Q(X_N, m)$$

where X_N is the observed data with N observations, and

$$m^s(\theta) \equiv m(Y_N^s(\theta)) = \arg\min_m Q(Y_N^s(\theta), m)$$

where Y_N^s is the data generated from the s-th simulation of the model given parameter vector θ. Then the indirect inference estimator of θ, $\hat{\theta}$, satisfies:

$$\hat{\theta} = \arg\min_\theta \left(\bar{m} - \frac{1}{NS}\sum_{s=1}^{NS} m^s(\theta)\right)' W_N \left(\bar{m} - \frac{1}{NS}\sum_{s=1}^{NS} m^s(\theta)\right) \tag{17}$$

where W_N is an arbitrary positive definite weighting matrix, and NS is the number of simulations. More details on the computation can be found in appendix section A.5.

6.5.1 Functional Form Assumptions and Parameters All of the parameters of the model can be seen in Table 6. Payments from central government are normalized to 1. The discount rate, β, is set to 97%. The death rate, γ, is set at 5.3%, which gives an average life-expectancy just under 19 years.[23] The cost of moving to a '0' star council is estimated in the model, but the relative cost of moving to the other councils is based on the relative cost of house prices and average local authority rent for council housing in the area. As expected, the '3' star council has the most expensive moving cost. The ten parameters that are estimated in the model are described in the panel (b) of Table 6.

6.5.2 Data Moments and Identification I match eleven moments in the estimation procedure. The moments are the distribution of the population across councils, the percentage of the population within each council each period who are migrating in, the percentage of moves that are to councils of the same quality, and the regression coefficients from the regressions of the distribution of pensioners across areas and the in-migration rate of areas on the star rating.

With the amount of moments in this model, a rigorous identification argument is difficult. Even still, the moments that have been chosen to be matched are informative about the parameters in model that are being estimated. Firstly, the distribution of pensioners across the councils and the percentage of the population within each council who migrate in is informative of mean quality of each council, the dispersion of individual qualities, the dispersion of initial beliefs, the dispersion of the noise in the social services experience, and the cost of moving. The percentage of moves that are to councils of the same quality is informative for the exogenous migration probability (as most moves to councils of the same quality are likely

[23]In 2005, life expectancy at age 65 for men in the England was 17.4 years and for women was 20.1 years (Office for National Statistics, 2013).

Description	Parameter	Value
Panel (a): Pre-set Parameters		
Central Government Payment (State Pension)	y^c	1.000
Discount Rate	β	0.970
Per Period Death Rate	γ	0.053
Relative Cost of Migrating to '1' Star Councils	k_1/k_0	1.005
Relative Cost of Migrating to '2' Star Councils	k_2/k_0	1.010
Relative Cost of Migrating to '3' Star Councils	k_3/k_0	1.110
Panel (b): Parameter Estimates		
Mean Quality of '0' Star Councils	μ_0	0.670
		(0.005)
Mean Quality of '1' Star Councils	μ_1	0.690
		(0.031)
Mean Quality of '2' Star Councils	μ_2	0.787
		(0.076)
Mean Quality of '3' Star Councils	μ_3	0.797
		(0.120)
Exogenous Migration Rate	δ	0.008
		(0.004)
Cost of Migrating to '0' Star Councils	k_0	2.98
		(0.219)
Dispersion of Initial Beliefs	κ	0.701
		(0.119)
Dispersion of Individual Qualities	τ	0.301
		(0.094)
Dispersion of Social Service Noise	σ	1.650
		(0.323)
Weight Given to Info Shock in Own Area	Δ_p	0.239
		(0.099)
Weight Given to Info Shock in Other Areas	$\Delta_{\bar{p}}$	0.086
		(0.041)

Notes: The cost of moving to '1', '2', and '3' star councils are pre-set as a relative cost of moving to a '0' star council (which is an estimated parameter). The information shock weight parameters, Δ_p and $\Delta_{\bar{p}}$, are determined by equations (15) and (16). Indirect Inference is used to estimate the parameters. The estimation procedure chooses a vector of parameter values to minimize the distance between features observed in the data and those generated by simulated data from the model, as displayed in Table 7. Standard errors in parentheses are derived from numerical derivatives as described in appendix section A.5.1.

Table 6. Model Parameter Values, Pre-set (a) and Estimates (b)

to be exogenous). All of these moments only use data from years before the information shock occurred. Finally, the regression coefficients, which make use of all of the years of data, are informative for the weights given to the information shock when pensioners update their beliefs, and also to uniquely pin down the previous belief parameters.

6.5.3 Estimation of the Model Table 7 presents the observed and simulated moments from the model. The model does well in terms of matching most of the targeted moments in panel (a), and untargeted moments in panel (b). Figure A7 in the appendix displays a comparison of the (untargeted) kernel density of distribution of pensioners across areas in the data and model, and shows that the model fits the distribution well.

Panel (b) of Table 6 displays the parameter estimates from the simulated model. The estimation reveals some interesting results.

The estimated parameter for the exogenous migration rate is 0.8%, which given that the overall percentage of pensioners that move in a year is 1.5%, implies that around 55% of moves are exogenous "family shocks" and cannot be explained by the model. This is not surprising given that pensioners migrate for a variety of reasons other than social services. As expected, the dispersion of initial beliefs is large, and much larger than the dispersion of individual qualities.

The dispersion of the noise in the observed social services parameter is large, meaning that learning is slow. In fact, given the size of the parameter estimate it would take an average pensioner living in an around area 20 years for their beliefs to get within 25% of their true quality. Given that the average life expectancy in the model is around 19 years, most pensioners have beliefs that are far from their true qualities. The large estimate for the dispersion of noise also means that pensioners can have both large positive and negative social services experiences. Large negative shocks contribute to the overall migration rate as after a bad shock pensioners will revise their beliefs downwards and be more likely to migrate.

The utility and cost parameters also reveal interesting results. Recall that in the model central government payments have been normalized to 1. For my sample years, the standard state pension from UK government is £5,400 a year (in 2015 £s). Therefore, for social service users in the model, the mean utility from a '0' star council is £3,620. The mean utility of '1' star councils is higher at £3,730. The mean quality of '2' and '3' star councils are both higher again at £4,250 and £4,300, respectively. The cost of moving to a '0' star council is £16,200, which is a substantial cost but may seem small when compared with estimates found in other migration models, such as Kennan and Walker (2011). However, when you consider than £16,200 is around twice the average yearly utility in the model these estimates are not that different than the estimates in Kennan and Walker (2011). It is also important to remember

Moment	Data	Model
Panel (a): Targeted		
% Population in Councils with '0' Stars	6.84	6.19
	(0.05)	
% Population in Councils with '1' Stars	49.64	51.03
	(0.04)	
% Population in Councils with '2' Stars	38.93	37.05
	(0.08)	
% Population in Councils with '3' Stars	4.59	5.72
	(0.01)	
Out Migration Rate in Councils with '0' Stars	1.69	1.61
	(0.95)	
Out Migration Rate in Councils with '1' Stars	1.84	1.67
	(0.84)	
Out Migration Rate in Councils with '2' Stars	1.42	1.65
	(0.62)	
Out Migration Rate in Councils with '3' Stars	1.97	1.88
	(1.02)	
% Migration to Council with Same Stars	39.84	39.61
	(0.10)	
Coefficient from main specification (β_D)	0.01	0.01
	(0.003)	
Coefficient from in-migration specification ($\tilde{\beta}_{D,in}$)	-0.0001	0.0002
	(0.0001)	
Panel (b): Untargeted		
% Population That Move In Year	1.50	1.50
	(0.05)	
In Migration Rate in Councils with '0' Stars	1.23	1.51
	(0.58)	
In Migration Rate in Councils with '1' Stars	1.46	1.69
	(0.55)	
In Migration Rate in Councils with '2' Stars	1.40	1.65
	(0.56)	
In Migration Rate in Councils with '3' Stars	1.39	1.87
	(0.42)	
Coefficient from out-migration specification ($\tilde{\beta}_{D,out}$)	-0.0003	-0.0004
	(0.0001)	

Notes: This table displays the targeted and untargeted moments from the model compared with that from the data. Standard errors for the data moments are displayed in parentheses.

Table 7. Data and Model Simulated Moments

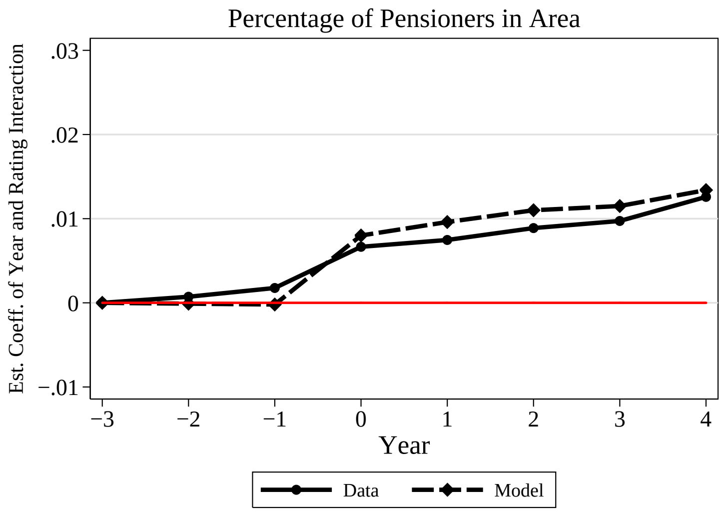


Figure 4. Visual Display of Impact of SSPR Star Ratings, Data versus Model
Notes: This graph plots the same coefficients for the data as displayed in Figure 3. The model estimates are calculated in the same way as the data estimates. While the baseline regression coefficient from Table 2 is matched in the model, the individual estimates displayed in this figure are untargeted.

that compared to other migration models, this paper focuses on a very different population, and the migration costs were found using different sources of variation. The high moving costs are a result of the weak relationship in my data between migration and area quality (according to their SSPR star ratings). Although in the model part of that observed weak relationship can be accounted for by pensioners having incorrect beliefs. Therefore, a model with perfect information or more accurate beliefs would estimate even higher moving costs.

	Average Difference in Utility After Information Shock	
	Present Discount Value	Yearly Flows
All Elderly Individuals	380	25
Those Influenced by Shock	6,900	600
Those Entering Model After Shock		
All	1,600	100
Those Influenced by Shock	41,000	2,800
Average Yearly Utility Before Information Shock: 8,750		

Notes: This table displays the difference in utility between pensioners in the model with an information shock and a counterfactual version of the model without an information shock. Those Influenced by Shock are pensioners in the model who acted differently in the model with an information shock than in the counterfactual version of the model without an information shock, i.e. the information shock caused them to move, not move, or changed where they move. Present Discount Value is calculated from the time of the information shock until the death of the pensioner, using the discount rate in Table 6. Those Entering Model After Shock are pensioners who enter the model in the three years immediately after the information shock, and gain the information shock as they enter the model. The Average Yearly Flow is the average yearly difference in utility (not discounted) in the ten years following the shock. All values are in 2015 £s.

Table 8. Welfare Analysis of Information Shock

6.5.4 Introducing an Information Shock As can be seen at the bottom of Table 6, to match the regression estimates from the data, pensioners have to place a 24% weight on the new information regarding their own area versus their prior beliefs, and only a 9% weight on information regarding the other areas versus their prior beliefs. The suggests that pensioners were much better informed about the star ratings in their own areas compared to others.

Figure 4 compares the regression estimates displayed in Figure 3 to simulations from the model. As can the seen in the figure, the model does well in replicating the introduction of the SSPR ratings effect on the distribution of pensioners across areas. At the time of the shock, '0' and '1' star councils see a drop in their population size, and '2' and '3' star councils see a rise. This shows that the information shock led to pensioners making better migration decisions, with the best councils seeing an increase in net migration and the worst councils seeing a decrease.

Table 8 presents a welfare analysis of the information shock, by comparing elderly individuals in the model after the information shock with themselves in a version of the model without an information shock. The table displays that the information shock led to increases in welfare, worth about £25 per year for the average pensioner, or a present discount value at the time of the shock of £380. This can also be seen in Figure A8. The reason that this amount is relatively small is due to not many pensioners moving each period and therefore for the majority of pensioners the information shock had no effect on their actions. In fact, less than 2% of pensioners in the model change an action as a result of the information release. Table 8 shows that those influenced by the information release - pensioners whose choice of whether or not to migrate, and where to migrate were affected - saw increases in welfare worth about £600 per year.[24] This is worth an equivalent present discount value at the time of the shock of £6,900. The real winners are those who enter the model after the information shock. Those who enter the model after the shock and are also influenced by the shock, experience increases worth around £2,800 per year. This is because any initial moves they make are better informed, and therefore they move to better matches, quicker. These pensioners also have their full lifespan in the model to benefit from the better matches and make fewer moves, which makes the present discount value (from the time they enter the model) of the information shock worth around £41,000.

6.6 Counterfactuals

6.6.1 Personalized Information Shock The SSPR ratings release did not provide personalized information to pensioners. This was modeled in the estimation of the model by the information shock bringing beliefs closer to the true mean quality for councils, not the true individual-specific values. Figure A9 displays shocks of different sizes that instead bring beliefs closer to the *true* qualities for every council, with the same weight given to the information for every council.[25] This mimics an information shock that gives individuals personalized information about their match qualities for every area. As can be seen in Figure A9 in the appendix, this type of information shock leads to different effects on migration and utility than the previous, not personalized, information shock. In each case, giving pensioners the personalized information shock results in large increases in the percentage of pensioners who migration in a year at the time of the shock, but the effect is temporary as pensioners immediately relocate to their preferred areas at the time of the shock and then remain there. This results in a temporary decrease in per period utility (as pensioners pay the large moving costs) but then average utility increases over time as more pensioners are in councils with which they have a good match.

[24]These results are consistent with Jin and Sorensen (2006) who study consumer choice of health plans after an information release regarding the quality of the plans. They find only a small fraction of individual decisions were affected by the release, but the affected individuals realized substantial utility gains.

[25]Specifically, at the time of the shock, equation (15) is instead $p_{i,j,t} = p_{i,j,t-1} + \Delta_{\tilde{p}} \left(q_{i,j} - p_{i,j,t-1} \right)$ for every council j.

6.6.2 Should Councils be Worried About Race-to-the-Bottom? A concern about making individuals more aware of the differences in a quality between areas is that everyone may migrate to the best areas. There are two main reasons why race-to-the-bottom is not a major concern in this model. First, pensioners have individual match qualities with areas, and therefore some pensioners will always prefer the '0' and '1' star areas. Second, moving costs are large and often well more than exceed any increases in utility that pensioners would gain by moving.

While I do not model the (supply) response of the councils, and do not deal with congestion effects, even in this model large information shock or changes in council quality do not result in large changes in the distribution of pensioners across areas. For example, if a single council changes its quality from '0' stars to '3' stars, the pensioner population grows by between 1.5%-2% according to whether pensioners are not informed about the change or whether they are given updated information on the change.

Even giving pensioners a perfect personalized information shock only increases the number of located in '2' and '3' star councils by 2% after 10 years. Therefore councils should not be worried about a large influx of welfare migrants following information releases, and hence do not have to engage in race-to-the-bottom.

7 Conclusion

While there is a substantial literature on welfare-induced internal migration, most studies find either no effect or a very modest one. This paper seeks to establish the role of information as part of the explanation for the lack of welfare-based migration. I assess whether the low observed rate of internal welfare migrants is due to individuals not knowing the quality of welfare programs in their area. I do this by studying how individuals respond to information releases regarding the quality of local public services.

I focus on pensioners in England, where social services provision is decentralized to a local level and displays a wide variation in quality. Using a policy called the Social Services Performance Review (SSPR), where the national government gave a publicly-released star rating of each area's social services on a scale from zero to three stars, I analyze the distribution of pensioners across areas before and after the information shock occurred. The hypothesis is that if information does play a role then, after the information is made public, we would expect areas that receive a higher star rating to see an increase in pensioners compared to the areas which performed poorly in the ratings. In my empirical strategy I find that a one increase in the publicly-released star rating leads to a 0.01 percentage points increase in the percentage of all pensioners that live in the area relative to others, which corresponds to a

1.3% increase. I interpret this increase as being caused by migration and provide evidence from migration data. The migration estimates suggest that the release of the SSPR does not appear to have had much of an effect on in-migration, but did impact out-migration, with pensioners less likely to leave areas that performed well in the ratings.

I use the results from my empirical investigation to motivate and estimate a search model with learning. The model suggests that there is a lot of noise in the learning process, and that it takes pensioners a long time to learn the true quality of social services in their area. To generate a response of the same magnitude as that observed in the data in response to the SSPR release, pensioners would need to place more weight on information released about their current area than about information regarding other areas. The best areas saw increases in net migration whereas the worst saw decreases in net migration. On average, this had a small positive effect on overall welfare, worth about £25 per year, which persists into the future. The main people who benefit are those induced to move, or not move, in response to the information shock, who experience welfare increases worth on average £600 per year. The form of the information shock, and whether it provides personalized information or not, is shown to be important, with personalized information releases potentially offering much greater welfare gains.

This paper shows that information can offer a partial explanation for why there is a lack of internal welfare-migration. The findings suggest that information releases - especially regarding the quality of local services or benefits - can be met with migratory responses and lead to greater differences in population compositions across areas.

References

Audit Commission (1995). *Local Authority Performance Indicators: Vol. 1, Education, Social Services and Total Expenditure.*

BBC (2002). *Social Services Ratings.* http://news.bbc.co.uk/hi/english/static/in_depth/health/2002/social_services_ratings/default.stm [Accessed: 2017/11/20].

Beaulieu, N. D. (2002). Quality information and consumer health plan choices. *Journal of health economics*, 21(1):43–63.

Bettinger, E. P., Long, B. T., Oreopoulos, P., and Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the h&r block fafsa experiment. *The Quarterly Journal of Economics*, 127(3):1205–1242.

Brookshire, D. S., Thayer, M. A., Tschirhart, J., and Schulze, W. D. (1985). A test of the expected utility model: evidence from earthquake risks. *Journal of political economy*, 93(2):369–389.

Brueckner, J. K. (2000). Welfare reform and the race to the bottom: Theory and evidence. *Southern Economic Journal*, pages 505–525.

Carrillo, P., Cellini, S. R., and Green, R. K. (2013). School quality and information disclosure: Evidence from the housing market. *Economic Inquiry*, 51(3):1809–1828.

Chen, Y. and Rosenthal, S. S. (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics*, 64(3):519–537.

Chernew, M., Gowrisankaran, G., and Scanlon, D. P. (2008). Learning and the value of information: Evidence from health plan report cards. *Journal of Econometrics*, 144(1):156–174.

Conlon, J. J., Pilossoph, L., Wiswall, M., and Zafar, B. (2018). Labor market search with imperfect information and learning. Technical report, National Bureau of Economic Research.

Conway, K. S. and Houtenville, A. J. (1998). Do the elderly vote with their feet?'. *Public Choice*, 97(4):663–685.

Conway, K. S. and Houtenville, A. J. (2001). Elderly migration and state fiscal policy: evidence from the 1990 census migration flows. *National Tax Journal*, pages 103–123.

Crépon, B., Ferracci, M., Jolivet, G., and van den Berg, G. J. (2018). Information shocks and the empirical evaluation of training programs during unemployment spells. *Journal of Applied Econometrics*, 33(4):594–616.

Dafny, L. and Dranove, D. (2008). Do report cards tell consumers anything they don't already know? the case of medicare hmos. *The Rand journal of economics*, 39(3):790–821.

Decreuse, B. and Tarasonis, L. (2016). Statistical discrimination in a search equilibrium model: Racial wage and employment disparities in the us.

Dennett, A. R. (2010). *Understanding internal migration in Britain at the start of the 21st century*. University of Leeds.

Department for Education (2018). *Compare school and college performance*. https://www.compare-school-performance.service.gov.uk/ [Accessed: 2018/05/19].

Dowding, K. (2005). *Local Government and its Discontents: Citizens' Preferences for Local Services: ESRC Full Research Report, R000223439*.

Dowding, K. and Mergoupis, T. (2003). Fragmentation, fiscal mobility, and efficiency. *Journal of Politics*, 65(4):1190–1207.

Dranove, D., Kessler, D., McClellan, M., and Satterthwaite, M. (2003). Is more information better? the effects of 'report card' on health care providers. *Journal of political Economy*, 111(3):555–588.

Eeckhout, J. and Weng, X. (2010). Assortative learning. *WPU Pennsylvania*.

Emmerson, C. and Leicester, A. (2002). *A survey of the UK benefit system*. Institute for Fiscal Studies.

Fernandez, J.-L. and Forder, J. (2015). Local variability in long-term care services: Local autonomy, exogenous influences and policy spillovers. *Health economics*, 24(S1):146–157.

Fernández, J.-L. and Snell, T. (2012). Survey of fair access to care services:(facs) assessment criteria among local authorities in england.

Figlio, D. N. and Lucas, M. E. (2004). What's in a grade? school report cards and the housing market. *The American Economic Review*, 94(3):591–604.

Fiva, J. H. and Kirkebøen, L. J. (2011). Information shocks and the dynamics of the housing market. *The Scandinavian Journal of Economics*, 113(3):525–552.

Friesen, J., Javdani, M., Smith, J., and Woodcock, S. (2012). How do school 'report cards' affect school choice decisions? *Canadian Journal of Economics*, 45(2):784–807.

Gelbach, J. B. (2004). Migration, the life cycle, and state benefits: How low is the bottom? *Journal of Political Economy*, 112(5):1091–1130.

Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of applied econometrics*, 8(S1).

Graves, P. E. and Knapp, T. A. (1988). Mobility behavior of the elderly. *Journal of Urban Economics*, 24(1):1–8.

Haisken-DeNew, J. P., Hasan, S., Jha, N., and Sinning, M. (2016). Unawareness and selective disclosure: The effect of school quality information on property prices.

Hussain, I. (2017). Consumer response to school quality information shocks: Evidence from the housing market and parents' school choices. mimeo.

Imberman, S. A. and Lovenheim, M. F. (2016). Does the market value value-added? evidence from housing prices after a public release of school and teacher value-added. *Journal of Urban Economics*, 91:104–121.

Jin, G. Z. and Sorensen, A. T. (2006). Information and consumer choice: the value of publicized health plan ratings. *Journal of health economics*, 25(2):248–275.

Jolivet, G., Jullien, B., and Postel-Vinay, F. (2016). Reputation and prices on the e-market: Evidence from a major french platform. *International Journal of Industrial Organization*, 45:59–75.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1):972–990.

Kennan, J. and Walker, J. R. (2010). Wages, welfare benefits and migration. *Journal of Econometrics*, 156(1):229–238.

Kennan, J. and Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251.

Leicester, A. and Shaw, J. (2003). *A survey of the UK benefit system*. Institute for Fiscal Studies.

Li, F. and Weng, X. (2017). Efficient learning and job turnover in the labor market. *International Economic Review*, 58(3):727–750.

Liebman, J. B. and Luttmer, E. F. (2015). Would people behave differently if they better understood social security? evidence from a field experiment. *American Economic Journal: Economic Policy*, 7(1):275–99.

Lockwood, B. and Porcelli, F. (2013). Incentive schemes for local government: Theory and evidence from comprehensive performance assessment in england. *American Economic Journal: Economic Policy*, 5(3):254–286.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.

Mastrobuoni, G. (2011). The role of information for retirement behavior: Evidence based on the stepwise introduction of the social security statement. *Journal of Public Economics*, 95(7-8):913–925.

Matikka, T. and Paukkeri, T. (2016). Does information increase the take-up of social benefits? evidence from a new benefit program.

McKinnish, T. (2007). Welfare-induced migration at state borders: New evidence from micro-data. *Journal of public Economics*, 91(3):437–450.

Meyer, B. D. et al. (1998). *Do the poor move to receive higher welfare benefits?* Institute for Policy Research, Northwestern University.

Moscarini, G. (2005). Job matching and the wage distribution. *Econometrica*, 73(2):481–516.

Moscone, F., Knapp, M., and Tosetti, E. (2007). Mental health expenditure in england: a spatial panel approach. *Journal of Health Economics*, 26(4):842–864.

NHS Executive (1995). *The NHS Performance Guide 1994-1995*.

Office for National Statistics (2001). *Census 2001: First results on population for England and Wales*.

Office for National Statistics (2013). *Life Expectancy at Birth and at Age 65 for Local Areas in England and Wales: 2009-11*.

Papageorgiou, T. (2014). Learning your comparative advantages. *Review of Economic Studies*, 81(3):1263–1295.

Papke, L. E. and Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1):121–133.

Pessino, C. (1991). Sequential migration theory and evidence from peru. *Journal of Development Economics*, 36(1):55–87.

Revelli, F. (2006). Performance rating and yardstick competition in social service provision. *Journal of Public Economics*, 90(3):459–475.

Santos, R., Gravelle, H., and Propper, C. (2016). Does quality affect patients' choice of doctor? evidence from england. *The Economic Journal*, 127(600):445–494.

Scanlon, D. P., Chernew, M., McLaughlin, C., and Solon, G. (2002). The impact of health plan report cards on managed care enrollment. *Journal of health economics*, 21(1):19–41.

Schwartz, A. L. and Sommers, B. D. (2014). Moving for medicaid? recent eligibility expansions did not induce migration from other states. *Health Affairs*, 33(1):88–94.

Smith, A. A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8(S1).

Social Services Inspectorate (2002). *Performance Ratings for Social Services in England 2001-2002*.

Wang, J., Hockenberry, J., Chou, S.-Y., and Yang, M. (2011). Do bad report cards have consequences? impacts of publicly reported provider quality information on the cabg market in pennsylvania. *Journal of health economics*, 30(2):392–407.

Wedig, G. J. and Tai-Seale, M. (2002). The effect of report cards on consumer choice in the health insurance market. *Journal of health economics*, 21(6):1031–1048.

Werner, R. M., Norton, E. C., Konetzka, R. T., and Polsky, D. (2012). Do consumers respond to publicly reported quality information? evidence from nursing homes. *Journal of health economics*, 31(1):50–61.

Wiswall, M. and Zafar, B. (2015). How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2):117–169.

A Appendix

A.1 Data Appendix

Data on the SSPR ratings come from the Department of Health's Social Services Inspectorate website.[26]

Local authority spending data and area characteristics come from the yearly Local Government Comparative Statistics (LGCS) available on the website of the Chartered Institute of Public Finance and Accountancy (CIPFA).

Data on state pension and other benefit claimants comes from the Department of Work and Pensions (DWP). A count of people claiming at least one key DWP working-age benefit by local authority is taken from the Work and Pensions Longitudinal Study (WPLS). Data on the pension-age client group comes from 5% samples of the DWP administrative data on the population over state pension age who were claiming at least one of the key benefits: attendance allowance, disability living allowance, incapacity benefit, pension credit, state pension, and severe disablement allowance.

Migration numbers into and out of local authorities come from Patient register data (1999-2008). This data is from National Health Service Central Register (NHSCR) and contains number of people moving from a GP in one area to another. It has both the inflow and outflow by broad age bands across all the areas in England and Wales. Dennett (2010) shows data is of high quality and the migration numbers match well with the 2001 Census.

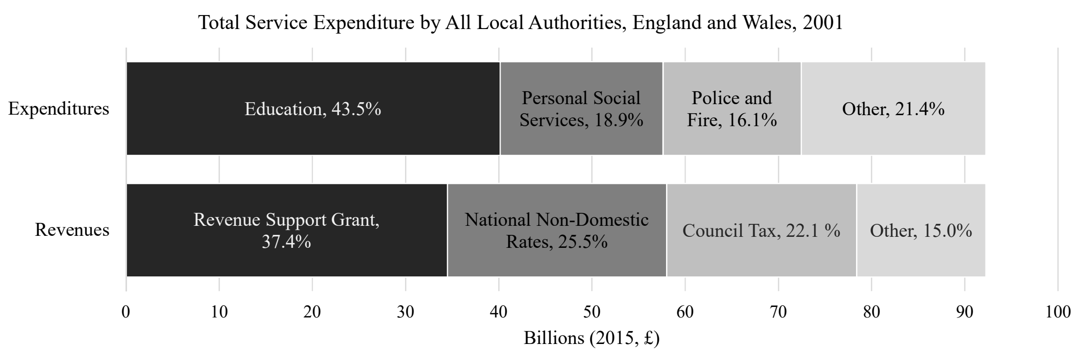
Various control variables are linked to the data by local authority, this includes house price data are from the Land Registry, and the average weekly full time wage comes from the Annual Survey of Hours and Earnings.

A.2 Constructing a Control Group (Predicting 2002 SSPR Ratings for Wales)

The 2002 SSPR Assessment included evidence from inspections and reviews, monitoring and performance indicators (Social Services Inspectorate, 2002). In order to ensure that the performance indicators had sufficient weight in the star rating system, and to provide an additional consistency check to ensure that councils were treated in the same way, a subset of 11 performance indicators were defined as the Key Performance Indicators, most of which exist for both England and Wales. In practice, a lot of weight was put on the assessors and not all of the same performance indicators are available for Wales.

To get around this I construct the control sample using a ordered logit regression model

[26]The Social Services Inspectorate was replaced by the Commission for Social Care Inspection in 2004 which was subsequently replace by the Care Quality Commission in 2009. Their website can still be accessed through the UK's National Archives website.



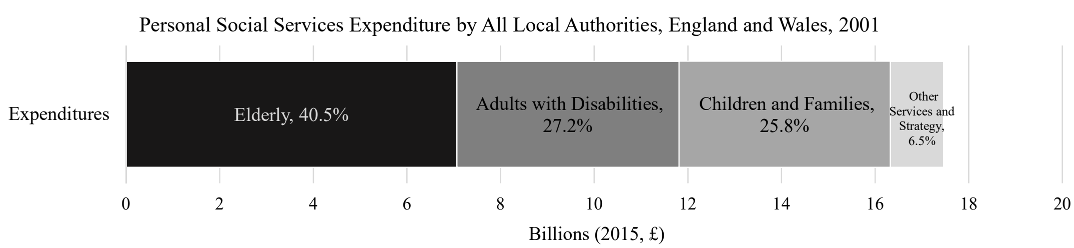


Figure A1. Revenues and Expenditures of All Local Authorities in England and Wales, 2001

Notes: The figure displays the total spending and revenue in 2001 of 148 local authorities in England and 22 local authorities in Wales. All amounts in 2015 £.
Source: Finance and General Statistics from the Chartered Institute of Public Finance and Accountancy (CIPFA).

Determinants of SSPR Ratings 2002		
	Ordered Logit	
Variables	Estimate	Std. Error
Older people helped to live at home (% over 65s)	0.0536***	(0.0163)
Adults and older people receiving a statement of needs (% population)	0.0419*	(0.0236)
Number of Day Care Clients		
Elderly	0.0022	(0.0012)
Physical Disability	-0.0035	(0.0047)
Learning Disability	0.0038**	(0.0033)
Mental Disability	0.0055***	(0.0035)
Accounting cost controls (Gross Total Cost)	✓	
Pseudo R^2	0.53	
Observations	148	

Notes: The sample includes all English Local Authorities (LAs) in 2002. The dependent variable is the 2002 SSPR star rating for each English LA. The accounting cost controls include the gross total cost in each service area related to PSS. Coefficients are displayed as odds ratios.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table A1. Predicting SSPR for Wales

to predict the 2002 SSPR ratings based on a wide range of personal social services accounting level data and comparable performance indicators (that exist for both England and Wales) for each local authority. I also include dummies for the accounting area of each local authority (to reflect that costs vary by area, as set out by the SSPR guidelines) and set Wales to be in the same cost group as the Midlands. The results from the ordered logit model can be seen in Table A1. Figure A2 displays the geographic variation in the predicted social service ratings.

It is worth noting that the ratings were based heavily on the performance indicators, and therefore were not driven by other area characteristics. This can be seen in Table A2, where another ordered logit model is displayed with the 2002 SSPR ratings regressed on a large number of area characteristics from 2001. These area characteristics do a very poor job of predicting the rating that a council would receive, and what is more, out of 28 variables only one is statistically significant above the 10% level.

Area Characteristics	Estimate	Std. Error
Percentage of Population		
Young (Aged 0-18)	-0.277	(0.189)
Elderly	-0.129	(0.122)
Working Age Benefit Claimants	-0.253	(0.335)
Unemployment Benefit Claimants	0.004	(0.699)
White	-0.021	(0.038)
Percentage of Elderly		
Disabled	0.090	(0.097)
Low Income	-0.075	(0.174)
Out-Migration	-0.896	(0.808)
In-Migration	0.670	(0.706)
Percentage of WA Benefit Claimants		
Disabled	-0.171	(0.205)
Low Income	-0.263	(0.209)
Council Type		
London Boroughs (Inner)	3.514**	(1.764)
London Boroughs (Outer)	1.052	(1.240)
English Metr. Districts	0.243	(0.691)
English Unit. Authorities	0.065	(0.676)
Conservative Dummy	0.506	(0.611)
Lib. Dem. Dummy	-0.335	(0.863)
Other Party Dummy	-0.714	(2.523)
No Overall Control Dummy	-0.310	(0.469)
Other Area Characteristics		
Crime Rate: Homicide	-17.430	(18.110)
Crime Rate: Burglary (Domestic)	-0.014	(0.010)
Crime Rate: Burglary (Non-domestic)	-0.095	(0.109)
Crime Rate: Vehicle Crimes	0.061	(0.049)
Average School Quality	-0.054	(0.042)
House Prices (log)	-0.360	(1.381)
Average Wages (log)	-0.703	(2.465)
Average Council Tax (log)	-1.915	(1.721)
Gross Social Services Spending (log)	0.256	(1.737)
Pseudo R^2	0.12	
Observations	148	

Notes: Omitted category for council types is English Counties, and for Political Control is the Labour Party. Social service spending is per capita of population.

Table A2. Ordered Logit Regression of 2002 SSPR Rating on Area Characteristics

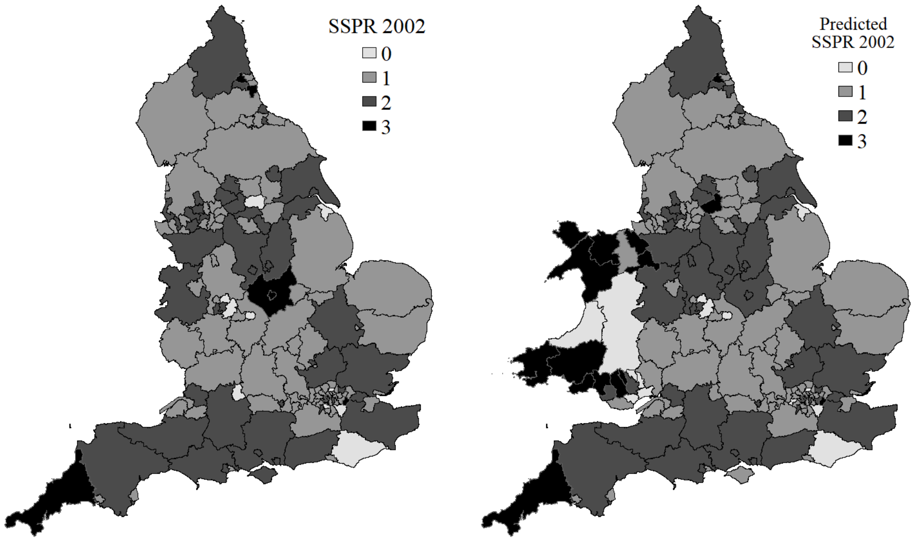


Figure A2. Geographic Variation: 2002 Actual and Predicted SSPR Ratings
Notes: This figure displays the actual 2002 SSPR star ratings in England (left) versus the predicted 2002 SSPR star ratings for England and Wales (right).

A.3 Testing the Identification Strategy

In this subsection I present some comparison statistics for England and Wales and test the identification strategy. Table A3 displays summary statistics for all local authorities in England and Wales, separated by their predicted SSPR. Figure A3 allows for visual inspections of the pre-reform years for all the main outcome variables. Table A4 tests the common trends assumption as described in section 4.

A.4 Alternate Specification

The specification in equation 1 assumes a linear effect in the star rating received. In this section I estimate an alternate specification which relaxes this assumption. In particular, I estimate

$$Y_{it} = \tilde{\beta}_0 + \sum_{r=0}^{3} \boldsymbol{\tilde{\beta}_{1,r} \tilde{R}_{r,i}} + \tilde{\beta}_2 P_t + \sum_{r=0}^{3} \boldsymbol{\tilde{\beta}_{D,r} (\tilde{R}_{r,i} \times P_t)} + \gamma' X_{it} + u_i + \tau_t + \epsilon_{it} \tag{18}$$

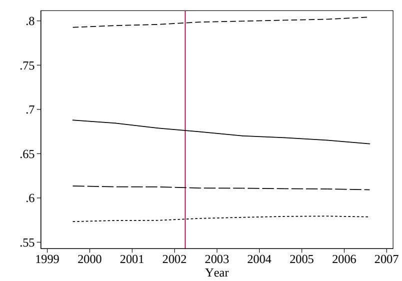
where the only change from equation equation 1 is that $\tilde{R}_{r,i}$ are dummy variables taking the value 1 area i received a 2002 SSPR rating of r, where $r \in \{0, .., 3\}$. The coefficients of interest are $\tilde{\beta}_{D,r}$.

Similar to before we can include Welsh councils as a control group and the specification is as follows,

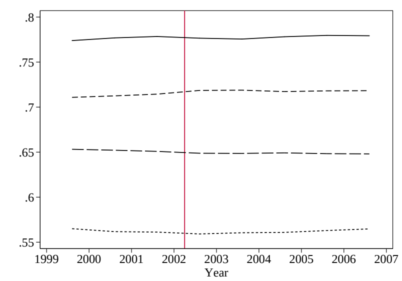
	Predicted 2002 SSPR Rating							
	England				Wales			
	0	1	2	3	0	1	2	3
Outcome variables								
Avg. Percentage of Pensioners	0.696	0.552	0.765	0.637	0.344	0.267	0.277	0.174
Pensioners	66,171	52,495	72,743	60,560	32,740	25,425	26,317	16,567
Population Aged Over 60	75,700	60,146	82,835	72,320	37,380	29,163	30,017	19,633
Total Population	387,214	297,184	387,778	339,080	172,740	128,662	128,683	81,500
Migration (Pop. Aged Over 60)								
In-Migration Rate (%)	1.447	1.443	1.357	1.578	1.974	1.784	1.383	2.033
Out-Migration Rate (%)	2.381	1.819	1.410	2.023	1.545	1.557	1.087	1.519
Net Migration Rate (%)	-0.934	-0.376	-0.0530	-0.445	0.429	0.227	0.296	0.514
Average House Price (2015 £)	172,819	129,314	106,938	230,960	92,042	75,358	76,115	74,608
Area Characteristics								
Average Weekly Income (2015 £)	1017	971	911	1048	841	837	814	808
School Quality, GCSEs	43.86	45.73	45.05	47.64	48.95	49.01	47.29	50.85
Weather, Rainfall	67.31	66.44	66.50	68.96	97.13	94.19	107.60	107.70
Weather, Average Temperature	10.59	10.24	9.949	10.80	9.668	10.09	9.840	9.481
Political Control of LA								
Labour Dummy	0.571	0.388	0.510	0	0.600	0.500	0.833	0.333
Conservative Dummy	0.143	0.188	0.255	0.800	0	0.125	0	0
Lib. Dem. Dummy	0	0.0824	0	0	0	0	0	0
Other Party Dummy	0	0.0118	0	0	0.200	0.375	0	0.667
No Overall Control Dummy	0.286	0.329	0.235	0.200	0.200	0	0.167	0
Average Council Tax	1135	1010	987	929	783	816	792	821
Gross PSS Spending on the Elderly	184	187	189	195	169	192	223	224
Gross PSS Spending on Other	303	243	215	252	219	227	233	198
Number of Local Authorities	7	85	51	5	5	8	6	3

Table A3. 2001 Mean Local Authority-Level Characteristics by Predicted 2002 SSPR Rating, England and Wales

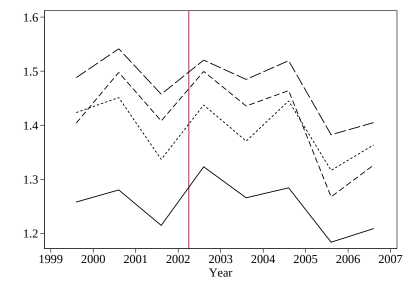
Notes: This table displays the same statistics as listed in Table 1, but local authorities are separated by predicted 2002 SSPR star ratings.



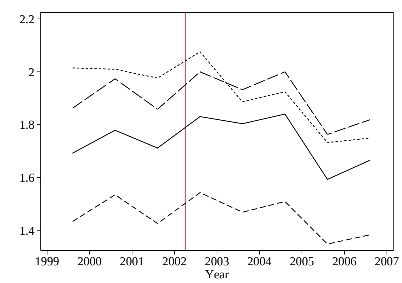
(a) % of Pensioners in Area



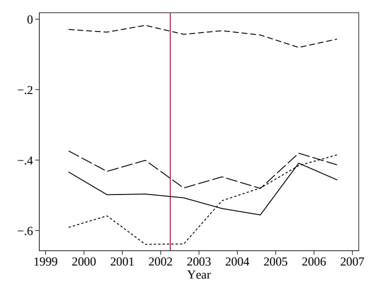
(b) % WA Benefit Claimants in Area



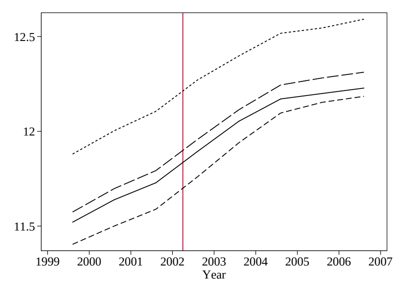
(c) In-Migration Rate (%)



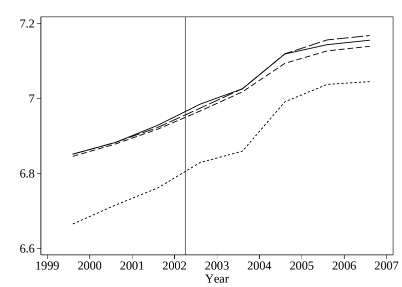
(d) Out-Migration Rate (%)



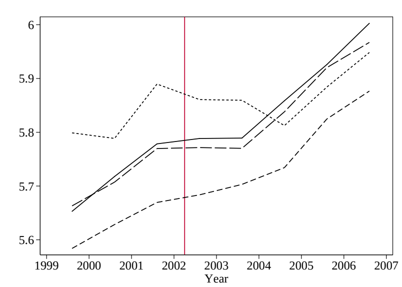
(e) Net Migration Rate (%)



(f) House Prices (log)



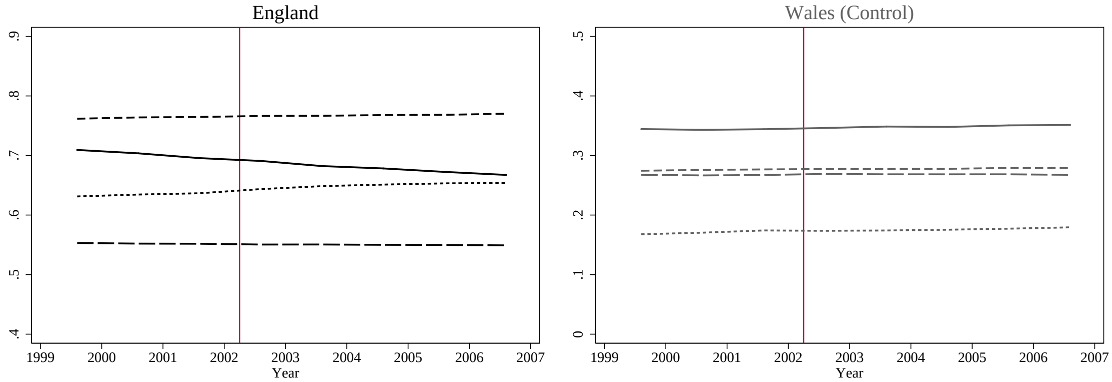
(g) Average Council Tax Bill (log)



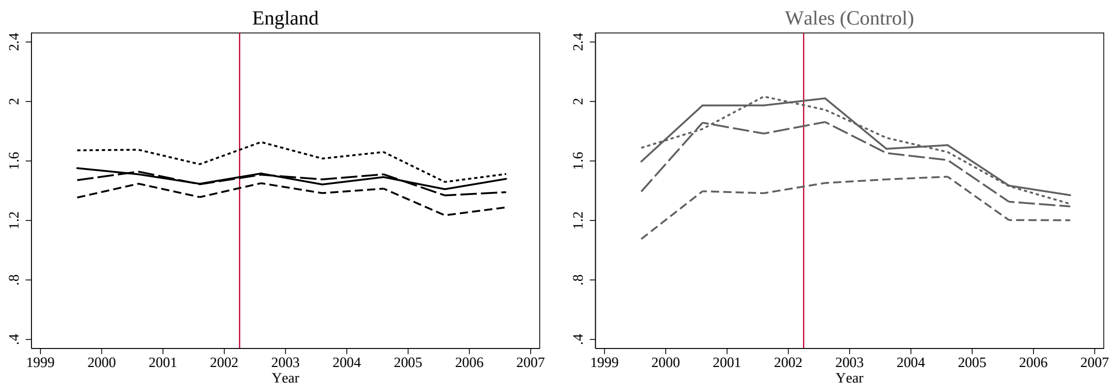
(h) PSS Spending (per capita, log)

Figure A3. Mean Trends, by 2002 SSPR rating

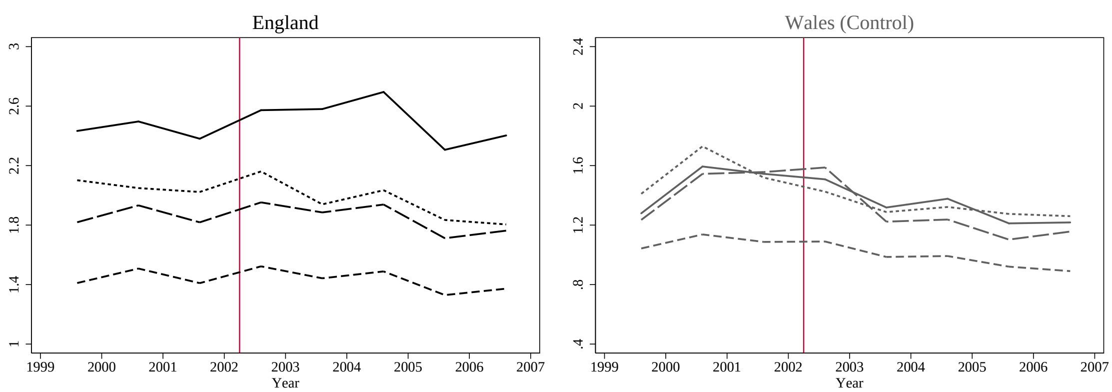
Notes: This figure displays the mean trends over time of all the main outcome variables for local authorities separated by 2002 SSPR rating. Subfigure (a) displays the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. Subfigure (b) displays number of Working-Age (WA) benefit claimants living in a LA as a percentage of all the WA benefit claimants in the country. The in-migration rate in subfigure (c) is calculated as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in that area at the beginning of the year. The out-migration rate in subfigure (d) is calculated similarly for those who migrate out of an area. The net migration rate in subfigure (e) is the in-migration rate minus the out-migration rate.



(a) % of Pensioners in Area



(b) In-Migration Rate (%)



(c) Out-Migration Rate (%)

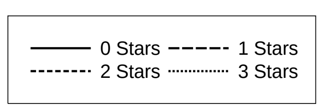


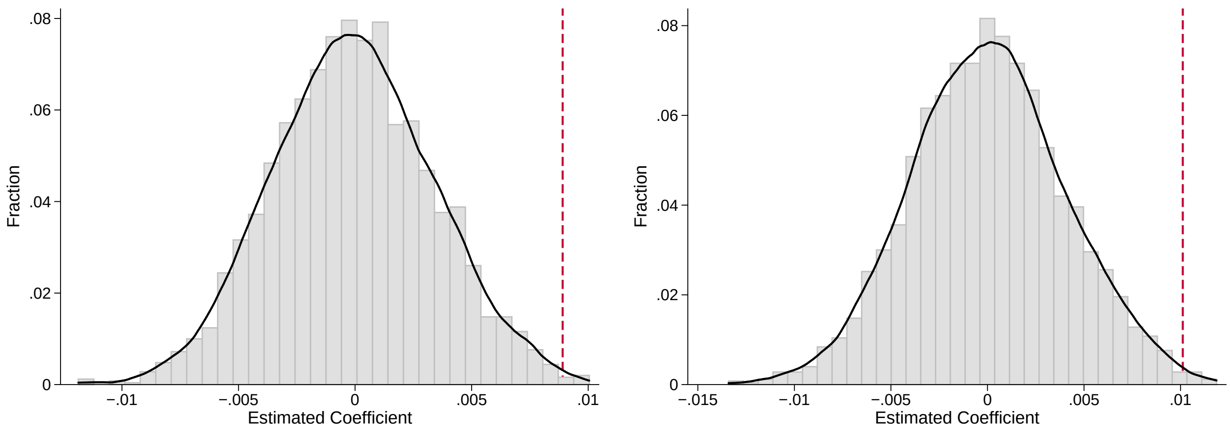
Figure A4. Mean Trends, by Predicted 2002 SSPR rating

Notes: This figure displays the mean trends over time of all the main outcome variables for local authorities separated by 2002 SSPR rating, for both England and Wales.

Variables	Testing Eq. (3) (1) p-value	Testing Eq. (4) (2) p-value
Percentage of Pensioners in Area	0.412	0.205
Percentage of Working Age		
Benefit Claimants in Area	0.765	0.623
Low Income	0.157	0.864
Disabled	0.865	0.602
Percentage of JSA Claimants in Area	0.634	0.682
Movers (Population Aged 60+)		
In-Migration Rate	0.659	0.736
Out-Migration Rate	0.359	0.276
Net Migration Rate	0.262	0.327
Movers (Population Aged 60-74)		
In-Migration Rate	0.770	0.765
Out-Migration Rate	0.238	0.175
Net Migration Rate	0.295	0.168
Movers (Population Aged 75+)		
In-Migration Rate	0.453	0.552
Out-Migration Rate	0.373	0.397
Net Migration Rate	0.207	0.261
House prices (Log)	0.480	0.922

Notes: Probability of rejecting the null hypothesis of similar time path between local authorities (LAs) with different SSPR Ratings in 2002 in the pre-treatment period when the null is true. P-values related to the null hypothesis $H_0 : \theta_{99} = \theta_{00} = \theta_{01} = 0$ from a regression of equation 3.

Table A4. P-values related to the null hypothesis $H_0 : \theta_{99} = \theta_{00} = \theta_{01} = 0$



Notes: These graphs display histograms of the distribution of the coefficient on percentage of pensioners in each area obtained by re-estimating the models in equations (1) and (2) after randomly reshuffling the star ratings of the local authorities 2,499 times. The left graph displays the model from equation (1) and the right graph from equation (2). The red stippled vertical lines show the placement of the coefficients from the models using the true star ratings, as displayed in Table 2. Kernel density plots are also displayed. The kernel density functions are estimated with the Epanechnikov kernel and bandwidths of 0.0006 and 0.0007, respectively.

Figure A5. The Distribution of Coefficients From Random Placebo Tests

$$\begin{gathered} Y_{it} = \tilde{\beta}_0 + \textstyle\sum_{r=0}^{3} \tilde{\boldsymbol{\beta}}_{\mathbf{1},\boldsymbol{r}} \tilde{\hat{\boldsymbol{R}}}_{\boldsymbol{r},\boldsymbol{i}} + \tilde{\beta}_2 P_t + \tilde{\beta}_4 (P_t \times E_i) \\ + \textstyle\sum_{r=0}^{3} \tilde{\boldsymbol{\beta}}_{\mathbf{5},\boldsymbol{r}} (\tilde{\hat{\boldsymbol{R}}}_{\boldsymbol{r},\boldsymbol{i}} \times \boldsymbol{P}_{\boldsymbol{t}}) + \textstyle\sum_{\boldsymbol{r}=\mathbf{0}}^{\mathbf{3}} \tilde{\boldsymbol{\beta}}_{\mathbf{6},\boldsymbol{r}} (\tilde{\hat{\boldsymbol{R}}}_{\boldsymbol{r},\boldsymbol{i}} \times \boldsymbol{E}_{\boldsymbol{i}}) \\ + \textstyle\sum_{r=0}^{3} \tilde{\boldsymbol{\beta}}_{\boldsymbol{DD},\boldsymbol{r}} (\tilde{\hat{\boldsymbol{R}}}_{\boldsymbol{r},\boldsymbol{i}} \times \boldsymbol{P}_{\boldsymbol{t}} \times \boldsymbol{E}_{\boldsymbol{i}}) \\ + \gamma' X_{it} + u_i + \tau_t + \epsilon_{it} \end{gathered} \tag{19}$$

Table A5 shows that the estimated effect of an area receiving a ‘0’ star rating is associated with a -0.02 percentage points decrease in the percentage of pensioners in those areas. As can be seen in the table, a rating of ‘3’ stars is estimated to be associated with 0.02 percentage points increase in the percentage of pensioners in those areas relative to ‘0’ star areas. The estimated effect of ‘1’ stars versus ‘0’ stars associated with a 0.01 percentage point increase in the population of pensioners, which is similar in magnitude to the estimated effect of ‘2’ stars versus ‘1’ stars. However, the estimated effect of ‘3’ stars versus ‘2’ is statistically insignificant and close to 0. This suggests that people do not view ‘2’ or ‘3’ stars as that different. However, it must be remembered that there are only a handful of councils that received 3 star ratings and therefore the estimates are not very precise.

	Dependent Variable: Percentage of Pensioners in Area	
		Incl. Welsh Control
Estimated effect of 3 stars versus 0 stars $(\widetilde{\beta}_{D,3} - \widetilde{\beta}_{D,0})$	0.0215** (0.00991)	0.0152 (0.0102)
Estimated effect of 2 stars versus 0 stars $(\widetilde{\beta}_{D,2} - \widetilde{\beta}_{D,0})$	0.0190*** (0.00658)	0.0146 (0.00962)
Estimated effect of 1 stars versus 0 stars $(\widetilde{\beta}_{D,1} - \widetilde{\beta}_{D,0})$	0.00987 (0.00623)	0.0184** (0.00922)
Estimated effect of 3 stars versus 2 stars $(\widetilde{\beta}_{D,3} - \widetilde{\beta}_{D,2})$	0.00251 (0.00910)	0.000636 (0.00804)
Estimated effect of 2 stars versus 1 stars $(\widetilde{\beta}_{D,2} - \widetilde{\beta}_{D,1})$	0.00916** (0.00448)	-0.00381 (0.00671)
Estimated effect of 0 stars $(\widetilde{\beta}_{DD,0})$	-0.0202** (0.00961)	-0.00803 (0.00799)
Local Authorities	148	170
Observations	1,184	1,360

Notes: The estimates in this table are from regression equation (18). The dependent variable is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. All regressions include LA fixed effects, year effects, and full controls. Regressions are for 170 English and Welsh LAs. Control variables are those listed under *Area Characteristics* in Table 1. Standard errors are clustered at the LA level.
* p < 0.1, ** < 0.05, *** p < 0.01.

Table A5. Alternative specification

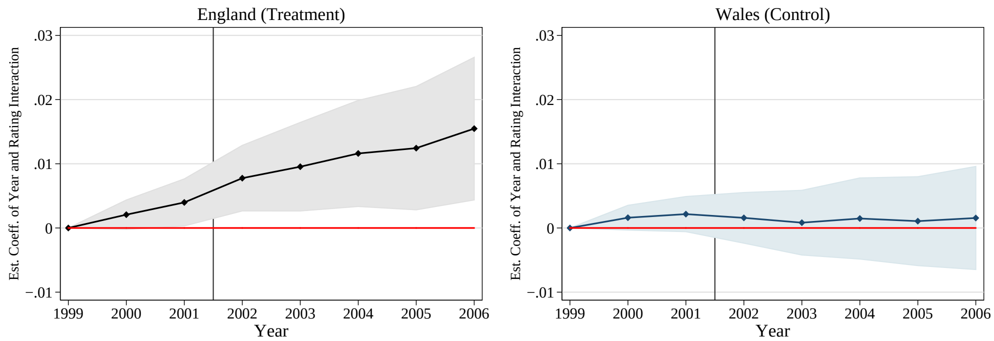


Figure A6. Visual Display of Impact of Predicted SSPR Star Ratings
Notes: This graph is a visual display of the estimates found in Table 2. This graph plots the β_{τ_t} coefficients for each year τ_t from the regression equation $Y_{it} = \beta_0 + \beta_1 \hat{R}_i + \sum_i \boldsymbol{\beta_{\tau_t}(\hat{R}_i \times \tau_t)} + \gamma' X_{it} + u_i + \tau_t + \epsilon_{it}$, where Y_{it} is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. Control variables are those listed under Area Characteristics in Table 1. The shaded gray area represents a 95% confidence interval around the estimates.

A.5 Computational Details

The stationary equilibrium of the quantitative model is found using value function iteration on equation (13), with moving costs included, to find the relevant policy functions. The model is estimated with six areas, and I set the number of points in the support of the belief distribution to six. As each pensioner has beliefs about every area, this means the relevant state space contains 46,656 elements for each area, and each of the six transition matrices contain $46,656^2$ elements. Increasing the number of points in the support of the belief distribution beyond six increases computational burden significantly and has very little effect on the estimates. Note that the discretization is used in order to calculate transition matrices and policy functions, but for the actual simulation there are no restrictions on beliefs (i.e. they do not lie on a grid). Therefore the discretization does not affect the calculation of per period utility, which is only influenced by decisions to migrate.

I nest the algorithm for solving the stationary equilibrium of the model into the following indirect inference estimation algorithm:

1. Choose an initial value of the parameter vector θ, θ_0, and set the objective function equal to infinity.

2. Solve for equilibrium of model using value function iteration.

3. Create $NS = 100$ data sets with my areas over 40 years with the information shock occurring at year 30. For each data set, compute the vector of moments $m^i(\theta)$

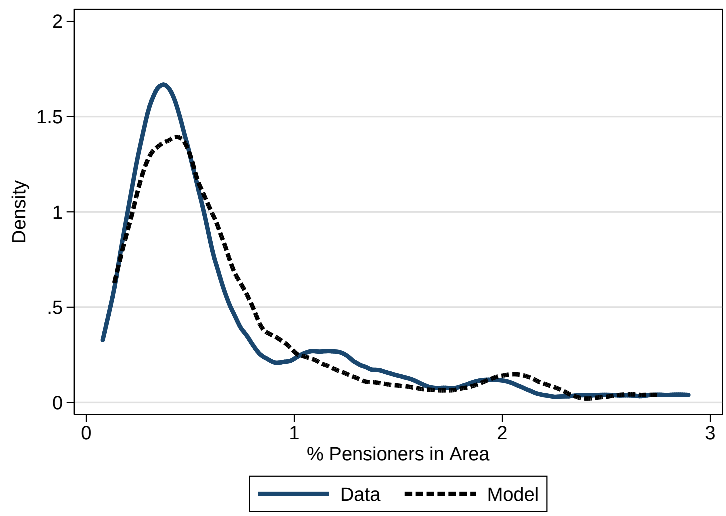


Figure A7. Kernel Density Function of Distribution of Pensioners Across 148 Areas
Notes: This figure displays a comparison of the kernel density function of the distribution of pensioners across all areas in the data versus the simulated model. This distribution is not targeted in the model.

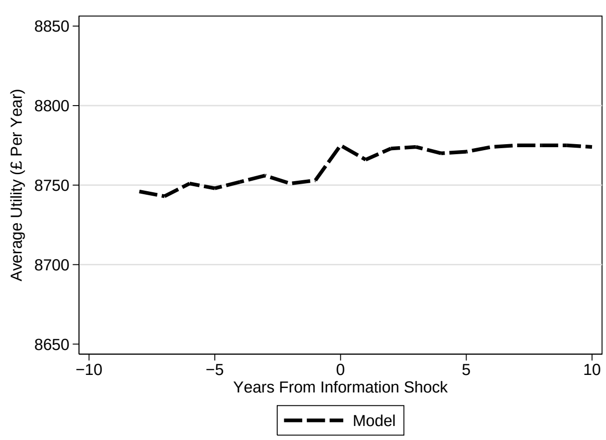


Figure A8. Average Utility (£ Per Year)

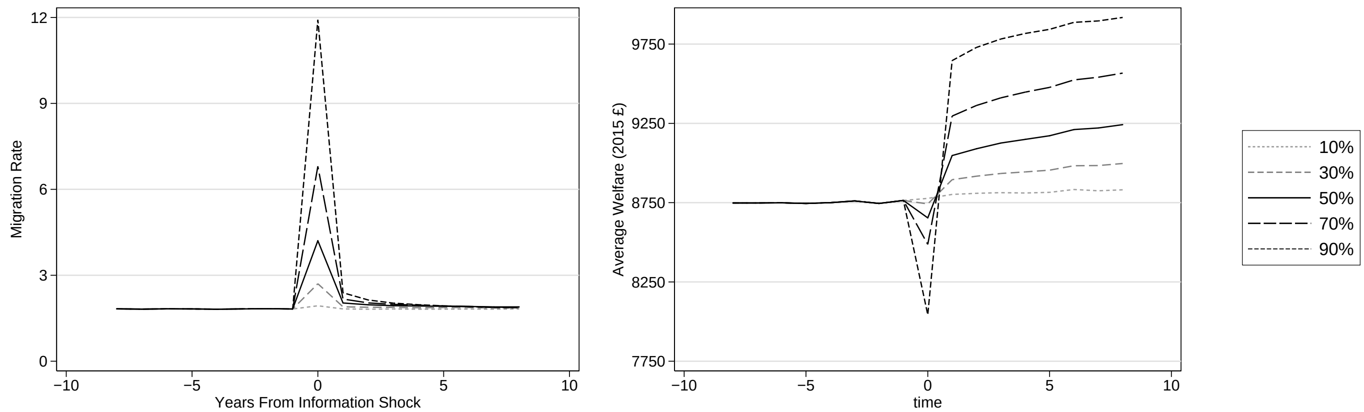


Figure A9. Change in Form of Information Shock (Personalized Information)
Notes: This figure displays the results from the simulated model using the parameter values listed in Table 6. The figure changes the form of the information shock to be a personalized shock as described in footnote 25, and displays the simulated response when the personalized information shock parameter, $\Delta_{\tilde{p}}$, takes different values.

4. Compute $\left(\bar{m} - \frac{1}{NS}\sum_{i=1}^{NS} m^i(\theta)\right)' \left(\bar{m} - \frac{1}{NS}\sum_{i=1}^{NS} m^i(\theta)\right)$ and update θ_{i+1} based on θ_i and the value for the objective obtained;

5. Repeat steps 2 through 4 as many times as possible.

I use a simulated annealing algorithm for the minimization of the objective and start with a number of initial guesses to ensure that the global minimum is attained.

A.5.1 Estimating Standard Errors The weighting matrix, W_N, in the indirect inference estimator, $\hat{\theta}$ (equation (17)) is the inverse of the variance-covariance matrix of the empirical moments, Ω. This is obtained by taking the variance-covariance matrices of the time average variables and regression coefficient from the DD regression, and constructing a block diagonal matrix. Then for a fixed NS, as the sample size tends to infinity, the indirect inference estimator is asymptotically normal with variance $var(\hat{\theta}) = \left(1 + \frac{1}{NS}\right) \left[\left(\frac{\partial m(\theta)}{\partial \theta}\right)' \Omega^{-1} \left(\frac{\partial m(\theta)}{\partial \theta}\right)\right]^{-1}$.