URBAN RESILIENCE AND SOCIAL SECURITY UPTAKE: NEW ZEALAND EVIDENCE FROM THE GLOBAL FINANCIAL CRISIS AND THE COVID-19 PANDEMIC

William Cochrane
School of Social Sciences, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand. Email: billc@waikato.ac.nz

Jacques Poot*
Te Ngira: Institute for Population Research, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand. Email: jacques.poot@waikato.ac.nz

Matthew Roskruge
School of Economics and Finance, Massey University, Private Bag 102904, North Shore Auckland 0745, New Zealand. Email: m.roskruge@massey.ac.nz

*Corresponding Author

ABSTRACT: This paper focuses on the spatial variation in the uptake of social security benefits following a large and detrimental exogenous shock. Specifically, we focus on the Global Financial Crisis (GFC) and on the onset of the COVID-19 pandemic. We construct a two-period panel of 66 Territorial Authorities (TAs) of New Zealand (NZ) observed in 2008-09 and 2019-20. We find that, despite the totally different nature of the two shocks, the initial increase in benefit uptake due to the COVID-19 pandemic was of a similar magnitude as that of the GFC, and the spatial pattern was also quite similar. We link the social security data with 146 indicator variables across 15 domains that were obtained from population censuses that were held about 2 years before the two periods. To identify urban characteristics that point to economic resilience, we formulate spatial panel regression models. Additionally, we use machine learning techniques. We find that the most resilient TAs had two years previously: (1) a low unemployment rate; and (2) a large public sector. Additionally, but with less predictive power, we find that TAs had a smaller increase in social security uptake after the shock when they had previously: (3) a high employment rate (or high female labour force participation rate); (4) a smaller proportion of the population stating ethnicities other than NZ European; (5) a smaller proportion of the population living in more deprived area units. We also find that interregional spillovers matter and that there are spatial clusters of resilient regions.

KEYWORDS: Urban economic resilience; social security; Global Financial Crisis; COVID-19; panel data; model selection; spatial econometrics; machine learning.

1. INTRODUCTION

Despite its remote location in the South Pacific, New Zealand (NZ) is tightly integrated with the global economy through trade, tourism, capital and migration flows, and strong digital connectivity (Plater and Claridge, 2000). Nonetheless, the country weathered the 2008 Global Financial Crisis (GFC) relatively well. More recently, effective public health and economic policies – including the strictest (but relatively short) lockdown measures among Organisation for Economic Co-operation and Development (OECD) countries – and the geographic advantage of being an island nation, muted in 2020 the adverse economic impact of the onset of the COVID-19 pandemic (Barrett and Poot, 2023; Gauld, 2023). Despite these favourable national outcomes, there are nonetheless large differences across people (Clyne and Smith, 2022) and places (Dyason *et al.*, 2021) in the impact of these sudden and large exogenous shocks that arrived from abroad.

In this paper, we focus on one important indicator of socio-economic impact, namely the increase in social security benefit uptake in the initial 6 months following each of the two shocks. This follows earlier work on determinants of social security benefit uptake in NZ labour market areas (Cochrane and Poot, 2009). Unlike other recent subnational-level work that has tended to focus on the COVID-19 shock and specifically on the effect of lockdowns, such as Bauer and Weber (2021), we pool data from the pandemic with those from the GFC. We construct a two-period (2008-09 and 2019-20) spatial panel of social security data across urban areas. We use data at the level of Territorial Authorities (TAs), which are the local government areas in NZ. The 66 TAs we distinguish can be considered to be local labour market areas (LMAs) since they contain mostly just one urban labour market and have little cross-boundary commuting.

We find that, despite the totally different nature of the two shocks, the initial increase in benefit uptake due to the COVID-19 pandemic was of a similar average magnitude as the increase due to the GFC. Moreover, the spatial pattern of the impact was also similar. This has been the case even though the initial policy responses to these shocks were entirely at the national level and, therefore, not spatially differentiated. Thus, there appear to be some stable underlying factors that determine the magnitude of the impact on the labour market, specifically in terms of job loss and/or income loss, that regions may experience following an unanticipated – and locally exogenous – detrimental shock arriving from abroad. When these factors operate similarly in the case of such distinct shocks, namely a financial markets disturbance and a public health threat, respectively, they may plausibly point to determinants of urban economic resilience, i.e. a certain level of resistance to a detrimental exogenous shock and a 'built-in' ability to recover relatively quickly.

During the last two decades, regional and urban economic resilience has become an important topic for understanding how economies at various spatial scales adjust to large exogenous shocks, although there have also been strong criticisms of the concept (Hassink, 2010). The literature makes it clear that there are a wide range of conceptualizations of economic resilience (Martin and Sunley, 2015). A common distinction is made between the 'engineering' perspective, in which a resilient system returns to the previous stable equilibrium after a shock and an 'ecological' perspective, in which the system moves to a new steady state (Groenewold, 2020; Modica and Reggiani, 2015). Martin and Sunley (2015) added to this the broader concept of 'adaptive resilience', which does not emphasise the long-run steady-state but instead focuses on the robustness of a complex system to exogenous shocks throughout paths of adjustment, either through 'built-in' mechanisms or through policy interventions (Hartal *et al.*, 2023).

A distinction is usually made between the initial 'resistance' phase following the shock and the subsequent 'recovery' phase. However, there is some evidence to suggest that LMAs may experience a long-run 'scarring' effect of an external shock-induced recession (Hershbein and Stuart, 2020). This is particularly the case in harder-hit metropolitan areas. Using Australian data, Andrews *et al.* (2020) found that these scarring effects of recessions are particularly present among young workers, who are, of course, a relatively large demographic group in metropolitan areas. Reviewing the international and Australian literature, Borland (2020) concluded that scarring effects are substantial. This suggests that the GFC may have had, through this scarring mechanism, a relatively lasting detrimental effect on regional resilience, which was still felt at the time of COVID-19. We report some evidence of this in the next section of the paper.

Faggian *et al.* (2018) argued that any empirical study of regional economic resilience should start with answering three fundamental questions: (1) "Resilience to what?"; (2) "Resilience of what?"; and (3) "Resilience over what period?". For the present paper, these questions have very specific answers. Firstly, we are investigating the resilience of NZ TAs to the onset of the GFC and COVID-19 shocks. Secondly, we are considering spatial variation in the extent to which the uptake of social security (and hence the associated public expenditure) deviated from the level observed in the relatively buoyant pre-shock period. Thirdly, we only focus on the initial impact by limiting the time frame to a 12-month period, with the initial shock occurring halfway through this period. This implies that we are specifically concerned with resistance and not with subsequent recovery.

The focus on the initial impact only is deliberate. We can plausibly argue that during the resistance phase the shock is totally exogenous and unanticipated. The study of determinants of resilience during the recovery phase must consider endogenous responses of firms and policymakers at national, regional and local levels (Hartel *et al.*, 2023). Other shocks may also emerge concurrently that make it difficult to define an endpoint for measurement of impact. For example, the final phase of the global COVID-19 pandemic and the start of the war in Ukraine overlap – which thwarts empirical assessment of the respective contributions of both events to the subsequent period of high inflation and depressed economic activity.

Using Italian data, Faggian *et al.* (2018) defined 'resistance' to the GFC as the growth in regional employment between 2007-08 and 2009-10, relative to national growth. 'Recovery' is defined as subsequent employment growth in 2011. They find considerable regional heterogeneity in both resistance and recovery. As is often the case in Italy, there is also a strong North-South divide, with the South being less resistant and having a slower recovery.

Additionally, Faggian *et al.* (2018) reconfirmed an earlier finding by Dijkstra *et al.* (2015) for all European regions that remote rural regions and large urban regions were more vulnerable to the GFC than intermediate urban and rural regions close to a city. The size of an urban area is an important predictor of vulnerability to the COVID-19 shock as well. Using Difference-in-differences (DID) analysis in the United States (US), Cho *et al.* (2021) found that employment rates decreased more in metropolitan areas than in non-metropolitan areas. Hamann *et al.* (2023) found similarly that in Germany the large cities were most affected. High employment density probably amplifies the effect of population density on COVID-19 infection rates.

The heterogeneity in the regional response to an exogenous shock is, not surprisingly, also related to the industry mix in the region. In the case of Italy, this has been confirmed by Rota *et al.* (2020). Using data on the US states, Kim *et al.* (2023) concluded that regional industrial structure is a strong determinant of the level of vulnerability of a region to unexpected recessionary shocks. Kim *et al.* found, for example, that essential industries with low personal interactions (such as non-store retailers and professionals working online) were the most resistant to the COVID-19 shock, while non-essential industries with high interpersonal interactions (such as tourism) were the most affected.

Using data from all 368 local authority districts in Great Britain, Houston (2020) stated that the pre 'lockdown' unemployment rate is an important predictor of the rise in unemployment in the first month of the lockdown at the onset of the COVID-19 pandemic. Pre-lockdown unemployment appears to matter more than the local industry mix. We shall see that this result holds in the NZ context also.

Whether the industry mix of a region is favourable or detrimental for weathering an exogenous shock would depend on the nature of the shock: consider, for example, the relatively large effect of COVID-19 on tourist destinations and that of the CFC on cities specialising in financial services. In general, we may expect that a diverse industry mix boosts regional resilience when the shock is strongly selective of certain industries. Using data from Ohio counties between 1997 and 2011, Brown and Greenbaum (2017) reported that counties with more diverse industry structures fared better during times of national employment shocks. Giannakis and Bruggeman (2017) stated that the dominance of manufacturing in a region in Europe lowered resilience to the GFC. Hundt and Grün (2022) reconfirmed this with data on German Spatial Planning Regions. Additionally, Hundt and Grün argued that regions with a greater share of public sector services are more resilient. We shall show that this is also the case in NZ.

NZ evidence on the impact of the COVID-19 pandemic on the labour market remains still relatively limited and has tended to be at the national level. Using their own survey (n=2002) designed to study life under the strict March-May 2020 nationwide lockdown, Fletcher *et al.* (2022) reported that this lockdown represented an unprecedented shock to the labour market. The national unemployment rate effectively doubled by week 3 of the lockdown. Particularly those on low incomes were affected, and close to 44 per cent of individuals lived in a household where at least one-member experienced job or income loss.

Clyne and Smith (2022) constructed an index of economic insecurity between 1999 and 2019 that reconfirmed the vulnerability of those on low incomes to the GFC shock and, by implication, to the COVID-19 shock. The indigenous Māori population and Pacific peoples more generally face the highest level of insecurity, but the Pākehā (i.e. the non-Polynesian population) faced the greatest percentage increase in insecurity following the GFC. Again, this analysis was only conducted at the national level.

The major public policy response to mitigate the detrimental effects of COVID-19 on the NZ labour market was the introduction of a Wage Subsidy Scheme (WSS) that was implemented in March 2020, and subsequently modified several times across five 'waves'. After an in-depth econometric investigation, Hyslop *et al.* (2023) concluded that the WSS was effective in that it particularly benefitted the most vulnerable firms and increased their survival rates. Additionally, positive employment effects for workers were also identified. However, most of this analysis was conducted at the national level, with sub-national results rather coarsely done by splitting of NZ into four regions: Auckland, Wellington, the rest of the North Island, and the South Island. After adjusting WSS take up rates in terms of differences between regional populations, Hyslop *et al.* (2023) concluded that "The ranking across regions is, however, similar whether raw or adjusted measures are used." (p.28). Further investigation of the WSS on benefit uptake rates is beyond the scope of the present paper.

The present paper is the first econometric analysis of the impact of the onset of the GFC and of the COVID-19 pandemic at the sub-national TA level. Given the likely impact on employment and income, we focus on social security benefit uptake as the indicator of impact, given that monthly data on this are readily available at the TA level, while the available survey data that inform on income and employment in NZ are quarterly and subject to relatively large sampling errors at this level of spatial disaggregation.

Internationally, this is also the first study to identify determinants of urban resilience following the GFC and COVID-19 shocks in one unified panel data setting. Brada *et al.* (2021) specified a spatial econometric model of relative employment change in 199 NUTS-3 regions in Central and Eastern Europe. While, like us, they considered regional resilience after the GFC as well as the COVID-19 shocks, their estimation is cross-sectional only – with data reflecting the regional resistance to, and subsequent recovery from, the GFC. The estimated coefficients were then used to simulate the likely impact of COVID-19 in the regions. In our case, we fully exploit the panel structure in the data and estimate the effects of the GFC and COVID-19 shocks simultaneously.

To uncover determinants of urban resilience, we link the social security data with 146 regional indicator variables across 15 domains that were obtained from population censuses that were held about 2 years before our specified GFC and COVID-19 observation windows. To identify urban characteristics that point to economic resilience, we are guided by stepwise model selection procedures (Lindsey and Sheather, 2010). For this, we first run the models with cross-sectional data in each of the two periods (2008-

09 and 2019-20). We then pool the two cross-sections to apply panel estimation techniques and account for spatial spillovers through designing spatial econometric models, broadly following the approach developed by Halleck Vega and Elhorst (2015). Finally, we use machine learning (ML) techniques implemented in Stata (Ahrens *et al.*, 2020), given that stepwise regression modelling can lead to the selection of over-fitted specifications (McNeish, 2015), to identify local predictors of resilience.

Our research reconfirms several of the findings in the earlier literature. We find that the most resilient TAs had about 2 years previously: (1) a low unemployment rate; and (2) a large public sector. Additionally, but with less predictive power, we find that TAs had a smaller increase in social security uptake after the shock when they had: (3) a high employment rate (or high female labour force participation rate); (4) a smaller proportion of the population stating ethnicities other than NZ European; (5) a smaller proportion of the population living in more deprived area units. We also find that interregional spatial spillovers matter. Similar to what Brada *et al.* (2021) found for Central and Eastern Europe, there tend to be also in New Zealand clusters of resilient regions.

The paper has five sections. Following this introduction, Section 2 describes the data and provides an exploratory analysis of determinants of the initial impact on social security uptake of the GFC and COVID-19 shocks by means of stepwise selection algorithms. Section 3 then reports on non-spatial and spatial panel models that are obtained after pooling the two cross-sections. Section 4 revisits the modelling by applying new machine learning techniques to the data. Finally, section 5 provides general conclusions and suggests avenues for further research.

2. DATA AND EXPLORATORY ANALYSIS

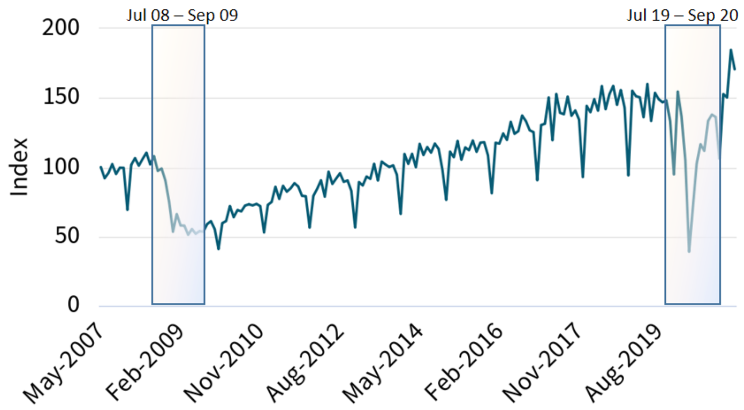
Data are predominantly drawn from NZ administrative sources and from the 2006 and 2018 population censuses. The spatial unit is the TA. Cochrane and Poot (2009) used 58 functional LMAs, based on travel to work data, to define the spatial unit of analysis. These LMAs mostly overlap with the TA regions used here. Rural populations are included in the TA data, but this has minimal impact on the data because NZ is highly urbanised (only 14 per cent of the NZ population lives in rural areas). Hence, we can interpret our geographical unit of analysis as being predominantly urban.

Before the amalgamation of Auckland TAs into one 'supercity' in 2010, data were available for 72 TAs (of which 7 made up the Auckland metropolitan area). Following the amalgamation, the TA database consists of Auckland and 65 other TAs. To obtain pre-2010 data for Auckland Super City, data from the seven TAs that made up Auckland were aggregated by means of population weights.

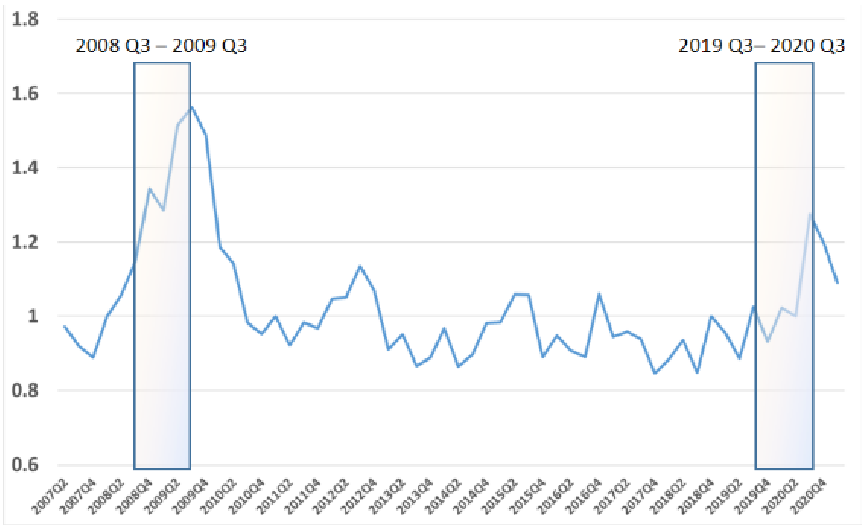
To define the time window for measuring the initial impact of the GFC and the COVID-19 pandemic on social security benefit uptake, no single time series of aggregate uptake is available, due to sweeping welfare reforms in 2013 that affected the types of social security available and the eligibility for these (Statistics New Zealand, 2022). Instead, we used two sources of high-frequency labour market indicators: the monthly online job advertisements index and the quarterly unemployment rate. Due to a range of factors, including the importance of the primary sector and tourism in the NZ labour market, high-frequency labour market and other economic indicators display strong seasonality. Fortunately, the initial impact of the two shocks was felt in roughly the same months in 2008 and 2020, respectively. Hence seasonality does not impact our estimation with the panel dataset that pools the two periods.

Figure 1 demonstrates that the appropriate timeframe for measuring the initial impact is to compare the third quarter (Q3) of 2009 with Q3 of 2008 for the GFC; and Q3 of 2020 with Q3 of 2019 for the COVID-19 pandemic. The monthly online job advertisements index declined from 101.4 to 52.4 between July 2008 and September 2009 and from 149.8 to 118.2 between July 2019 and September 2020 (Figure 1a). Job advertisements are always at their lowest during the December month. The effect of the strict lockdown from 25 March until 13 May 2020 (Alert Level 4 until 27 April, followed by Alert Level 3) is clear from the very low level of job advertisements in April and May 2020. Similarly, the unemployment rate divided by the unemployment rate four quarters previously peaked in Q3 2009 at 1.56 and in Q3 2020 at 1.28 (Figure 1b).

The dependent variable, *growth_ben*, measures the growth in social security benefit uptake. The data have been sourced from the Ministry of Social Development. For the GFC shock, *growth_ben* is defined as the sum of the average number of the four types of benefits (unemployment, sickness, domestic purposes and invalid) in the third quarter of 2009 in each TA minus the corresponding number in the third quarter of 2008, expressed as a percentage of the TA census usually resident population in 2006. Although the impact of the monetary policy responses to the GFC did differ across regions, quarterly unemployment data (Markham, 2020) show that the effect of the GFC on the labour market was felt in all regions from the third quarter of 2008 onward.



(a) Monthly Online Job Advertisements Index



(b) Unemployment Rate Divided by Unemployment Rate Four Quarters Previously

Figure 1. Defining the Window of Initial Impact by Labour Market Indicators. Source: https://www.stats.govt.nz/experimental/covid-19-data-portal and https://infoshare.stats.govt.nz

Following the 2013 Social Security (Benefit Categories and Work Focus) Amendment Act, the social security terminology and types of benefits have been changed. Consequently, for the COVID-19 shock, *growth_ben* is defined as the aggregated number in a TA in the categories: 'Jobseekers – Work Ready Benefit' (JS-WR); 'Jobseekers – Jobseeker Support – Health Condition and Disability'; and 'Benefit – Other' (average of months of July, August, and September 2020) minus the corresponding number in the third quarter of 2019, divided by the TA census usually resident population in 2018. Except for the Auckland lockdown in August-September 2020, there were mostly no regional differences in COVID-19 public health measures during our observation window.

Recent research has shown that the aggregated number of people receiving any kind of income-tested social security benefit is a more effective indicator of excess labour supply (and therefore of the short-run impact of the GFC and COVID-19) than JS-WR because the former is more highly correlated with the surveyed national unemployment rate than the latter (Rea and Maloney, 2021).

Table 1 provides descriptive statistics on the change in social security benefit uptake in a TA at the onset of either of the two shocks and a range of potential local determinants. Guided by the literature, we identified 15 domains of socioeconomic data that could potentially provide indicator variables that could predict resistance to the shocks, i.e. a relatively smaller increase in social security benefit uptake. Data on one pre-selected indicator in each of the 15 domains is reported in Table 1. A total of 146 indicators are available in the dataset. The selected domains capture population scale, age structure, ethnicity, openness, wealth, the elasticity of labour supply, human capital, public sector activity, casualisation of employment, social capital, labour market disequilibrium, industry structure, industry diversity, deprivation, and income.

The indicator variables are all sourced from the census previous to the shock considered, i.e. the 2006 census for the GFC shock and the 2018 census for the COVID-19 shock. The exception is the industry structure variable, which measures the expected total employment growth that would have occurred in the TA during the twelve months observation window (i.e. Q3 2008 to Q3 2009 for the GFC shock and Q3 2019 to Q3 2020 for the COVID-19 shock) if the TA industries grew at national industry growth rates. The source is the quarterly Household Labour Force Survey. This is also referred to as the Bartik index (Cochrane and Poot, 2020).

Table 1 shows that the initial impact of the GFC on social security benefit uptake was of a similar magnitude to that of the COVID-19 pandemic: a mean increase across TAs of 1.86 per cent versus 2.23 per cent respectively. The standard deviation was also similar in both cases (about 0.8).

Table 1. Variable Definitions and Descriptive Statistics. Sources: Described in the text.

Domain	**Indicator**	**Variable name**	**Variable definition**	**Period**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Correlation with growth_ben (1)**
Local labour market impact of global shock	Growth in social security benefit uptake	growth_ben	Growth in the total number of social security benefit recipients as a percentage of the total population	Q3 2008 - Q3 2009	1.86	0.86	0.33	4.31	1.00
				Q3 2019 - Q3 2020	2.21	0.83	0.78	5.54	1.00
1. Population scale	Territorial Authority population	lnpop	Natural logarithm of the census usually resident population	2006	9.98	1.04	7.95	13.79	0.39*
				2018	10.17	1.07	8.10	14.04	0.22
2. Age structure	Youth Dependency Ratio	youth_dep	Population aged 0-14 as a percentage of the population aged 15-64	2006	34.60	4.98	21.73	47.41	0.43*
				2018	32.19	4.15	22.10	42.77	0.37*
3. Ethnicity	Ethnic composition	pnoneuro	One hundred minus the percentage of population stating European ethnicity	2006	26.96	9.47	15.68	55.18	0.58*
				2018	18.25	11.19	6.17	50.46	0.48*
4. Openness	Geographic Mobility	geo_mob	Percentage of population who lived at a different address five years ago	2006	51.15	4.41	40.13	66.42	0.00
				2018	43.24	4.29	33.09	55.06	-0.26*
5. Wealth	Percentage in Rental Accommodation	prental	Percentage of households that rent the dwelling they occupy	2006	24.72	5.13	15.14	36.89	0.24
				2018	27.25	5.31	16.40	43.66	0.31*
6. Elasticity of labour supply	Female Labour Force Participation Rate	fem_lfpr	Those employed or unemployed and actively seeking work as a percentage of the population aged 15 and over	2006	62.64	4.43	53.26	76.62	-0.46*
				2018	63.47	4.56	53.78	78.25	-0.26*
7. Human capital	Percentage of population with tertiary education	ptertiary	Percentage of population aged 15 and over who had obtained a Bachelor degree or higher	2006	7.19	3.73	2.25	27.04	-0.11
				2018	12.95	5.19	5.46	36.96	-0.15
8. Public sector	Percentage public sector employment	pubsector_emp	Percentage of total employment by industry who are employed in the public sector	2006	14.80	4.61	6.15	28.24	-0.33*
				2018	14.19	4.78	6.28	28.75	-0.38*
9. Casualisation of employment	Percentage Self Employed	self_emp	Percentage of total employment whose employment status is self-employed	2006	23.47	6.42	7.89	36.86	0.03
				2018	18.61	4.76	8.00	30.52	0.03
10. Social capital	Percentage volunteering	Pvol	Percentage of the population aged 15 and over who volunteered for one hour or more per week	2006	16.38	2.27	11.30	21.60	-0.38*
				2018	14.66	1.91	10.06	19.59	-0.30*
11. Labour market disequilibrium	Unemployment rate	ue_r	Those unemployed and actively seeking work as a percentage of the labour force	2006	4.58	2.09	1.48	13.66	0.63*
				2018	5.65	2.67	0.96	17.77	0.49*
12. Industry structure	Projected Employment Change	pprjempch	Sum of regional industry shares times national industry employment growth during the year of the shock (Bartik index)	Q3 2008 - Q3 2009	-1.96	0.75	-3.28	0.35	-0.16
				Q3 2019 - Q3 2020	2.54	1.01	0.66	4.91	-0.12
13. Industry diversity	Industry diversity index	pdiversity_ind	One hundred minus one hundred times the sum of squared shares of industries in total employment	2006	89.37	3.33	79.04	92.57	0.29*
				2018	90.25	2.43	81.72	92.45	0.24
14. Deprivation	Prevalence of deprivation in deciles 9 & 10	pnzdep910	Percentage share of TA population in area units with deprivation index in deciles 9 and 10 nationally	2006	21.88	23.64	0.00	100.00	0.46*
				2018	24.85	25.49	0.00	100.00	0.57*
15. Income	Log of median income	lnmedpinc	Natural logarithm of median personal income	2006	10.03	0.12	9.75	10.39	-0.05
				2018	10.29	0.15	0.93	10.66	-0.33*

Notes: * after a correlation coefficient indicates significance at the 5% level or better.

Besides the standard descriptives, Table 1 also shows the correlation of each indicator variable with benefit uptake. This provides a first indication of which variables are likely to play a role as predictors of TA-level resistance to the GFC and COVID-19 shocks. Here, and in the subsequent regression analyses, observations are weighted – when the estimator allows it – by analytical weights that are proportional to TA population size.

Figure 2 compares the spatial distribution of the increase in social security benefit recipients between the GFC and COVID-19 shocks. The impact is mostly felt in the north and along the east coast of the North Island. The south of the South Island is much less affected. The maps also show spatial clustering of the effect of the shocks, which needs to be considered in the econometric modelling.

Figure 2 suggests that the correlation between TA benefit uptake growth during the GFC and the onset of COVID-19 is quite high (in fact, the simple correlation coefficient is 0.62). Using machine learning, to be formally introduced in section 4, it can be shown that GFC benefit growth is one of the selected predictors for estimating the impact of the onset of COVID-19 on benefit growth. This is consistent with the scarring effect discussed in the previous section. The effect is reinforced in TAs with a relatively high level of deprivation. In contrast, a large share of employment being in the public sector offsets the scarring effect to some extent.

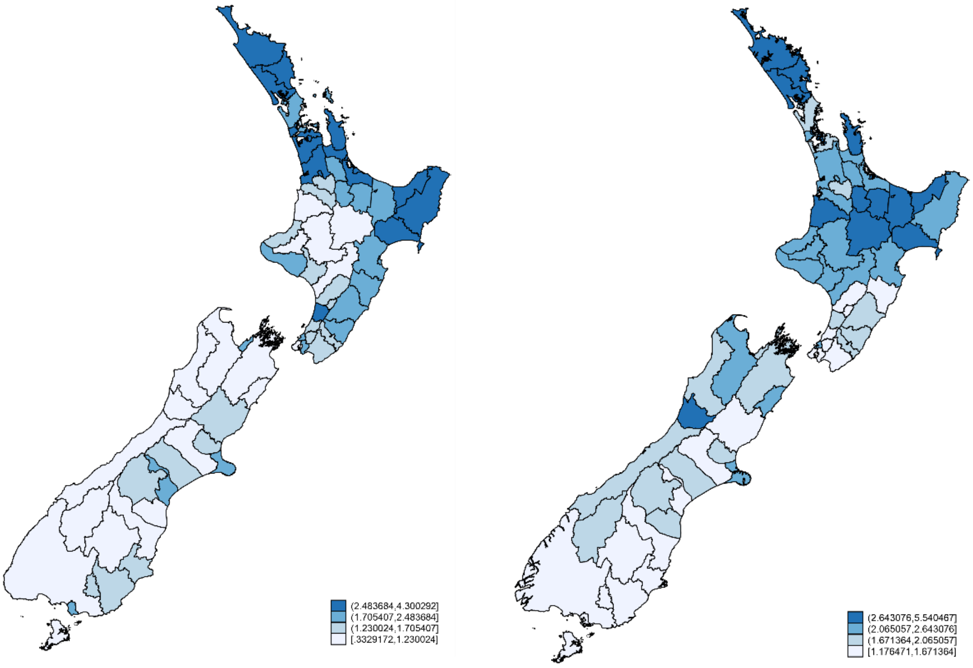
Following the GFC, benefit uptake is greater in the urban areas with larger populations (*lnpop*). This is not surprising since the initial impact of a large financial shock is mostly felt in metropolitan areas. The correlation between benefit uptake growth and population is not statistically significant at the 5 per cent level in the COVID-19 pandemic case.

Given that NZ provides a relatively generous old age pension from age 65 that is not income tested, the effect of a shock is more likely to be felt among those with young dependents, where social security support is less, and recipients must pass low-income and wealth tests. We measure age structure by *youth_dep*, the population aged 0-14 as a percentage of the population aged 15-64. This variable has a statistically significant correlation with benefit uptake after both shocks.

Another strong predictor at this descriptive level is ethnic composition. The indigenous Māori population and non-western migrant groups (particularly those from the Pacific Islands) have worse social and economic outcomes than other groups. Table 1 shows that TAs with a larger non-European population share (*pnoneuro*) experienced greater increases in benefit uptake after both shocks.

Geographic mobility is often considered an important mechanism for a local area to adjust to an exogenous shock. Table 1 shows that this does not appear to be the case at the time of the GFC. However, TAs where a large percentage of the population lived at a different NZ address five years previously (*geo_mob*) were less affected by the COVID-19 shock in terms of benefit uptake.

The percentage of households that rent the dwelling they occupy (*prental*) may be considered a proxy for a lack of wealth, given that equity in a dwelling is the main source of wealth for NZ households. TAs with a larger percentage of households renting were more affected by COVID-19.



Increase in TA Social Security Uptake (Q3 2009 Minus Q3 2008) of the 2006 TA Census Population

Increase in TA Social Security Uptake (Jul-Sep 2020 Average Minus Jul-Sep 2019 Average) as a Percentage of Census Population the 2018 TA

Figure 2. Social Security Benefit Recipients Increase: GFC vs. COVID-19. Source: Described in the text.

It is well known that the wage elasticity of labour supply is greater among females than among males. Consequently, we would expect TAs with a relatively large female labour supply to have a buffer against negative labour demand shocks. We find indeed that the TAs where the female labour force participation (*fem_lfpr*) was high experienced less of an impact on social security uptake. The correlation is statistically significant for both the GFC and COVID-19 shocks. In this context, it is notable that Klein *et al.* (2021) found gendered impacts of changing social security payments during COVID-19 lockdowns in Australia. On the other hand, the percentage of the labour force that is self-employed, *self_emp*, is not correlated with TA benefit uptake.

The level of human capital of the TA labour force (measured by *ptertiary*, the percentage of the population aged 15 and over who had obtained a Bachelor's degree or higher) had no statistically significant effect on the post-shock increase in benefit uptake. In contrast, the percentage of total employment by industry employed in the public sector (*pubsector_emp*) was, in terms of simple correlation, a strong predictor of which TAs were the least affected by the shocks in terms of benefit uptake.

The descriptive cross-sectional correlations also show a relationship with a common social capital variable: the percentage of the population aged 15 and over who volunteered for one hour or more per week (*pvol*). TAs with relatively high levels of social capital, as proxied by volunteering, experienced a lower increase in social security benefit uptake.

The strongest predictor of a post-shock increase in benefit uptake is the previous census unemployment rate (*ue_r*). The correlation of TA industry structure (measured by the sum of regional industry shares times national industry employment growth, *pprjempch*) with TA benefit uptake growth is, as expected, negative for both the GFC and the COVID-19 pandemic but not statistically significant. Industry diversity, measured by one hundred minus one hundred times the sum of shared shares of industries in total employment (*pdiversity_ind*), is only correlated with the benefit uptake increase after the GFC. Hence TAs with greater industry concentration were less affected by the GFC.

Socio-economic vulnerability in NZ is measured by a deprivation index that can be calculated at a fine spatial scale, such as a census area unit (Salmond *et al.*, 1998). We find that TAs in which the percentage of the population in area units with a deprivation index value in the 9^{th} or 10^{th} decile nationally (*pnzdep910*; i.e. they are the most deprived) are, as expected, also the TAs where the increase in benefit uptake following the two shocks was the greatest. Deprivation is a much stronger predictor of benefit uptake than TA median income. The negative correlation between benefit uptake and the natural logarithm of median personal income (*lnmedpinc*) is only statistically significant in the case of the COVID-19 pandemic.

Most of the 15 indicator variables that are listed in Table 1 are correlated with the cross-sectional variation in the growth in benefit uptake for at least one of the two shocks (*ptertiary*, *self_emp* and *pprjempch* are the exceptions) and plausible mechanisms can be suggested for the correlation in each case. Even with this small subset of 15 out of 146 indicators, there are potentially more than half a million regression models to consider. We use the leaps-and-bounds algorithm (Furnival and Wilson, 1974) implemented in Stata (command *vselect*) to identify the best regression model for each given number of regressors. The observations are weighted by the Census Usually Resident Population (CURP) of each TA. We select the most parsimonious model (i.e. with the least number of regressors) by means of the Bayesian Information Criterion (BIC), given that this criterion penalises most for additional regressors and that stepwise selection procedures tend to yield over-fitted models (Lindsey and Sheater, 2010). The results are shown in Table 2.

Using the BIC criterion, the optimal number of regressors (out of 15) in the case of the GFC data is four. The census unemployment rate *ue_r* is present in every step and is, therefore, the most robust predictor. We conclude that the TAs that were the most resistant to the onset of the GFC had the lowest unemployment rates two years previously.

Interestingly, in the case of COVID-19, the unemployment rate at the time of the previous census is also the strongest predictor of benefit take-up, except in the first step when the indicator of deprivation *pnzdep910* was selected. The optimal number of regressors for predicting social security benefit increase following the onset of COVID-19 is six.

Although the fit of the optimal models is equally good (with an R-squared of 0.709 and 0.738, respectively), the predictors differ but the unemployment rate and the rate of self-employment do feature in both optimal models. Hence, on balance, having a relatively large proportion of the workforce being self-employed is a sign of vulnerability rather than entrepreneurship (Blanchflower, 2004). Many of these self-employed are likely to be casual workers.

Social capital, measured by the percentage volunteering *pvol* and a favourable industry structure (*pprjempch*) did boost resilience after the GFC but were not predictors in the optimal model for COVID-19. In contrast, lower growth in benefit uptake after COVID was found in TAs where a larger share of the workforce was working in the public sector (*pubsector_emp*), where a smaller proportion of households rented their home (*prental*) and, interestingly, where industry diversity (*pdiversity-ind*) was less, i.e. industry concentration was greater. The negative coefficient of the logarithm of TA population in the COVID-19 regression indicates that the impact of the pandemic was greater in the larger urban areas, similar to what Hamann *et al.* (2023) found with German data.

Table 2. Classic Model Selection by Stepwise Regression. Source: Authors' Calculations.

		GFC			COVID-19	
Number of regressors selected		**BIC**	**Number of regressors selected**		**BIC**	
	1	98.4509		1	95.3238	
	2	67.79925		2	62.87758	
	3	64.53109		3	61.07219	
	4	**62.44913**		4	62.02884	
	5	63.93597		5	58.05612	
	6	66.24296		**6**	**53.4226**	
	7	69.8397		7	55.78845	
	8	73.4433		8	58.12506	
	9	77.32888		9	61.68205	
	10	81.29644		10	65.42391	
	11	85.25761		11	69.24009	
	12	89.2736		12	73.20069	
	13	93.39363		13	77.26886	
	14	97.50691		14	81.45107	
	15	101.6961		15	85.63061	
Variable	**Number of regressors when the variable is included**	**Coefficient when k=4**	**Robust std. err.**	**Number of regressors when the variable is included**	**Coefficient when k=6**	**Robust std. err.**
ue_r	1,2,3,4	0.382	0.034	2,3,4,5,6	0.218	0.030
pnzdep910				1		
pubsector_emp	2,3			2,3,4,5,6	-0.074	0.010
pdiversity_ind				3,5,6	0.140	0.028
pvol	3,4	-0.126	0.020			
self_emp	4	0.060	0.013	4,5,6	0.050	0.018
prental				4,5,6	0.049	0.011
pprjempch	4	-0.254	0.063			
lnpop				6	-0.135	0.046
constant		0.320	0.351		-11.417	
Number of obs		66			66	
R-squared		0.709			0.738	
Root MSE		0.359			0.296	

Notes: Each cell in the top half of the table refers to a separate regression, with the corresponding number of selected regressors. The regression method is weighted least squares, with analytical weights given by the population that was usually resident in the TA at the time of the census (CURP) preceding the crisis. Bold type rows indicate the number of regressors at which BIC is minimised and the corresponding BIC value.

While the simple descriptive analysis of this section has yielded some interesting similarities and differences between the onsets of GF and COVID-19 on TA-level resilience, there are three major deficiencies. The first is that as yet, we have not taken into account the panel structure of the data, i.e. repeated observations from the same TAs. Panel data estimators can account for unmeasured time-invariant features of TAs that may impact resilience.

Secondly, even though the initial GFC and COVID-19 shocks were national, the effects they have on TAs may lead to spatial spillovers. These two deficiencies will be addressed by the spatial panel data estimations that we report on in the next section. The third issue is whether the selected potential predictor of resilience for each of the 15 domains is the best among the variables that can be extracted from the available data sources. In Section 4, we will apply machine learning techniques to test the robustness of the patterns we observe in the selection of indicator variables.

3. PANEL DATA ESTIMATION

Considering that the census unemployment rate turned out to be the strongest predictor of benefit uptake in the descriptive analysis, we proceed with estimating a fixed effects (FE) panel model with a time trend. The TA data are weighted by the average population over the 2006-2018 period. The coefficient of *ue_r* with a panel FE estimator is 0.265, which is in between the values shown in Table 2 and statistically significant at the 5 per cent level (with robust standard errors). The time trend is not statistically significant. A Hausman test suggests that the random effects (RE) estimator is more efficient, but the RE and FE estimates are in fact quite similar (without the time trend, the RE estimate is 0.249 and the FE estimate is 0.208, both significant at the 1 per cent level). The Hausman test statistic is 0.44, which is not statistically significant (df = 2). Estimated without a time trend, this suggests that an increase in a TA's unemployment rate of 1 percentage point between 2006 (pre-GFC) and 2018 (pre-COVID) would imply a 0.21 to 0.25 percentage point increase in social security benefit uptake in the short-run following a national exogenous shock. Figure 3 shows the scatterplots of the data for the GFC and COVID-19 crises respectively. The size of the circles is proportional to the TA populations.

To identify additional variables that robustly enhance the RE panel model of benefit uptake we resort again to the *vselect* algorithm. This yielded *self_emp*, *pubsector_emp* and *pdiversity_ind* as important additional variables. The estimated coefficients of this RE panel model are reported in column (1) of Table 3. All variables in this panel model are statistically significant at the 5 per cent level or better. The coefficient of the unemployment rate increases to about 0.3. TAs with a relatively large share of the workforce being self-employed see a slightly greater increase in benefit uptake following an exogenous shock. On the other hand, a larger share of the workforce in public sector employment lowers the social security effect of the initial shock. Greater concentration of industry (i.e. lower *pdiversity_ind*) reduced the impact of a shock. This may seem surprising but the impact of regional specialization on social security uptake following a shock is theoretically ambiguous.

In the case of the COVID-19 shock, regions that specialised in tourism would have benefited from the government's WSS that provided income even if the businesses had a significant drop in revenue or were temporarily closed. In other TAs with a high concentration of certain sectors, firms' market power may have provided sufficient capital to weather the shocks; or demand was pre-dominantly export-oriented and, at least initially, less affected.

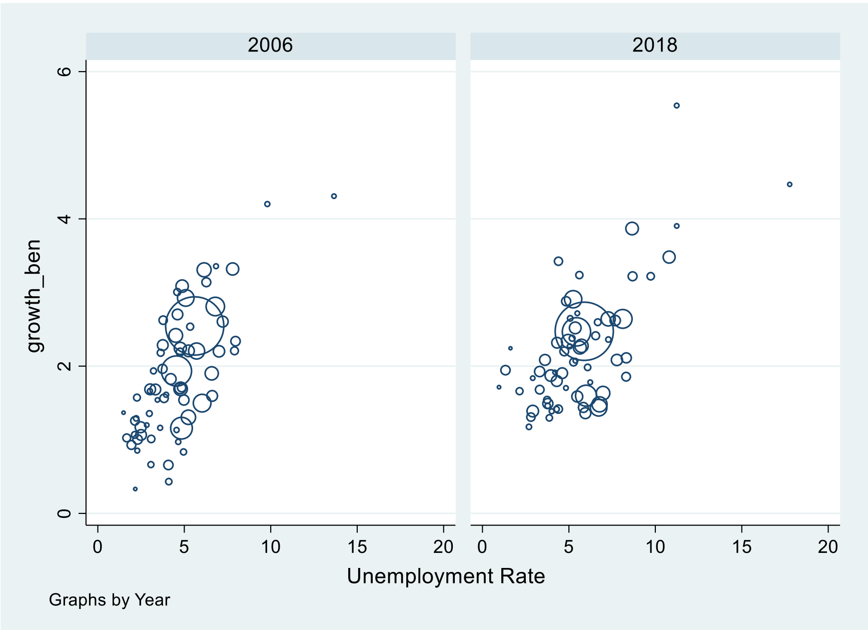


Figure 3. Pre-shock Unemployment Rates and Post-shock Social Security Uptake Increase. Source: Authors' Calculations.

The remainder of Table 3 reports the results of estimating a range of spatial econometric panel models (the Stata command is *spxtregress*). These models take account of spatial spillovers. Ignoring spatial spillovers may bias upward the effect of the included variables and also lead to lower estimated standard errors, i.e. yielding greater statistical significance than is actually the case.

The most general spatial model is the General Nested Spatial (GNS) model (Elhorst, 2014), which, in a panel setting, takes the following form:

$$\boldsymbol{y}_t = \rho \boldsymbol{W}\boldsymbol{y}_t + \alpha\boldsymbol{\iota}_N + \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\theta} + \boldsymbol{\mu} + \zeta_t\boldsymbol{\iota}_N + \boldsymbol{u}_t \quad (1a)$$

$$\boldsymbol{u}_t = \lambda\boldsymbol{W}\boldsymbol{u}_t + \boldsymbol{\varepsilon_t} \quad (1b)$$

where $\boldsymbol{y}_t$ is here a 66×1 vector consisting of one observation of benefit uptake increase for each TA at time t (t = 2008-09 or 2019-20); $\boldsymbol{\iota}_N$ is a unit vector associated with the constant term α; $\boldsymbol{X}_t$ is a 66×K matrix of K explanatory variables observed at time t and $\boldsymbol{\beta}$ is the associated K×1 parameter vector of effects of these variables on benefit uptake. The spatial weights matrix $\boldsymbol{W}$ is a positive 66 × 66 matrix which describes the structure of dependence between spatial units. In this application, the spatial weights are proportional to the reciprocal of the distance between pairs of TAs. The weights are normalised to add to one. $\boldsymbol{W}\boldsymbol{y}_t$ are the endogenous spatial spillover effects among the dependent variable, i.e. growth in benefit uptake, while $\boldsymbol{W}\boldsymbol{X}_t$ are the exogenous spillover effects of the independent variables across TAs.

The model also includes spatial and period-specific effects, $\boldsymbol{\mu}$ and $\zeta_t\boldsymbol{\iota}_N$ respectively. These may be treated as fixed effects or as random effects. However, given that we only have two periods (shocks) in our data, estimation with FE estimators is not likely to be precise. Hence, we will display RE estimates only. $\boldsymbol{W}\boldsymbol{u}_t$ represents the interaction effects among the disturbance terms of the different observations. The strength of spatial dependence between TAs is measured by the spatial diffusion parameters ρ and λ. Similarly, $\boldsymbol{\theta}$ is a K×1 vector of response coefficients that measure the average impact of variation in exogenous explanatory variables in surrounding areas.

The estimates of the most general case we consider are found in column (6) of Table 3. Using the notation of Equation 1(a) and 1(b), $\boldsymbol{\mu}$ represents random effects and, given that only two periods are considered, the time-fixed effect $\zeta_t\boldsymbol{\iota}_N$ has been deleted (the time dummy was insignificant in any case in the non-spatial panel model). Columns (2) to (5) and (7) represent models that result from applying restrictions to the model in column (6). Column (2) shows estimates of the spatial lag model which has $\boldsymbol{\theta}$ = 0 and λ = 0. Column (3) represents the spatial error model, in which

$\boldsymbol{\theta} = 0$ and $\rho = 0$. The estimates in column (4) are those of the Durbin spatial lag model, which has $\lambda = 0$. The estimates of the Durbin spatial error model with $\rho = 0$ are shown in Column (5). Finally, column (7) only allows for spatial lags of the exogenous variables and hence has $\lambda = \rho = 0$.

Table 3. Non-spatial and Spatial Random Effects Models. Source: Authors' Calculations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	**Non-spatial RE model**	**Spatial lag model**	**Spatial error model**	**Durbin spatial lag model**	**Durbin spatial error model**	**General spatial model**	**Spatial lagged X model**
ue_r	0.298***	0.249***	0.275***	0.234***	0.232***	.231***	.236***
self_emp	0.024**	0.020	0.022	0.008	0.008	0.008	0.008
pubsector_emp	-0.062***	-0.048***	-0.050***	-0.047***	-0.048***	-0.047***	-0.047***
pdiversity_ind	0.059***	0.040*	0.045*	0.035	0.036	0.036	0.038
constant	-4.415**	-3.681*	-3.113	-33.500**	-37.079**	-35.790199*	-36.572**
Spatially weighted							
growth_ben		0.538***		0.272		0.113	
e (growth_ben)			0.638**		0.406	0.365	
ue_r				0.045	0.154	0.117	0.127
self_emp				0.078*	0.087*	0.084*	0.086**
pubsector_emp				-0.0105*	-0.132**	-0.121	-0.130**
pdiversity_ind				0.343*	.384**	0.369*	0.378**
N	132	132	132	132	132	132	132
AIC		203.6	210.9	203.9	203.2	205.1	202.5
BIC		226.7	234	238.5	237.8	242.6	234.2
R-squared	0.635	0.653	0.633	0.680	0.682	0.681	0.682

Notes: * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Figure 2 already suggests the presence of spatial correlation. The spatial lag model (column (2)) and the spatial error model (column (3)) confirm this. The spatial correlation coefficients are 0.538 and 0.638 respectively. As expected, the coefficients of the explanatory variables are now smaller than those in the case of the non-spatial model (column (1)). Statistical significance is generally lower as well.

However, bringing in the spatial lags of the explanatory variables renders both ρ and λ (in columns (4) and (5) respectively) to be no longer statistically significant, but several of the spatial lags of the explanatory variables are statistically significant. The general spatial model is over-parameterised which can be seen from the relatively high values of the Akaike Information Criterion (AIC) and the Bayesian Information

Criterion (BIC). We conclude that the statistically best supported specification is that of the spatial lagged X (SLX) model of column (7); which is also the model that Halleck Vega and Elhorst (2015) consider the preferred model where there are no a priori theoretical considerations regarding which spatial model is appropriate. The SLX model suggests that there are two variables that robustly predict local resistance to the onset of an exogenous shock: a history of lower structural unemployment and the abundance of public sector jobs. The coefficients of these two variables are, as expected, a bit smaller than in the case of the non-spatial model: 0.236 and -0.047 respectively. However, the spatial model shows that the abundance of public sector jobs in surrounding regions is also beneficial (the coefficient is -0.130). Additionally, when the self-employed in surrounding TAs lose employment following a shock, benefit uptake in the TA at the centre increases (the coefficient is 0.086).

Finally, the spatial effect of industry diversity is interesting. More diverse industries in surrounding TAs (higher values of *pdiversity_ind*) increase the social security impact of the shock on the region of interest (the coefficient is 0.378).

4. MACHINE LEARNING APPROACHES

Up to this point, we have been considering a set of 15 specific variables, one for each of the 15 domains of socio-economic indicators, that were motivated by the literature to date. However, we have available 146 indicators in each of the 66 TAs that we observe twice (once for the onset of the GFC shock and once for the onset of the COVID-19 shock). In principle, we could repeat the analysis of the previous section with various alternative subsets of variables. This process has the danger of generating a set of predictors that fit the available data very well but may not yield accurate predictions in the case of other shocks. Since the objective of our paper is to identify indicators at the TA level that will predict socio-economic resilience to a future, as yet unspecified, global shock, regression methods that penalise both overfitting and the bias introduced by omitting relevant variables are expected to have superior performance.

In recent years machine learning (ML) techniques have been developed that can provide a robust set of predictors among a very large set of potential predictors. Molina and Garip (2019) provide a short introduction to these new developments in the social sciences. A review for economists is given by Mullainathan and Spiess (2017). The subset of ML techniques that is appropriate in the present context is that of Supervised Machine Learning (SML), where training data on inputs $\boldsymbol{X}$ (in our case characteristics of TAs) are linked to a desirable outcome $\boldsymbol{y}$ (i.e. low social security benefit uptake after a shock) with the goal of learning what function of $\boldsymbol{X}$ would give the best prediction of $\boldsymbol{y}$ once a new set of data on $\boldsymbol{X}$ is obtained.

Essentially SML accepts a trade-off between bias and variance by minimising

$$\left(\boldsymbol{y} - f(\boldsymbol{X})\right)'\left(\boldsymbol{y} - f(\boldsymbol{X})\right) + \pi R(f) \qquad (2)$$

in which the left-hand side of (2) reduces to the residual sum of squared errors in Ordinary Least Squares (OLS) regression when $f(\boldsymbol{X}) = \boldsymbol{X\beta}$. The right-hand side of (2) is called the regulariser, which penalises functions that generate variance in predictions. The weight π can be thought of as the relative price of variance. In OLS that price is zero, but a function f is then created in which some strong predictors of $\boldsymbol{y}$ in the sample data are given too much influence in prediction out of sample. The least absolute shrinkage and selection operator (LASSO) adds a regulariser that equals the sum of the absolute value of the estimated parameters of f (e.g. Hastie et al. 2015). Hence, if $f(\boldsymbol{X}) = \boldsymbol{X\beta}$ and each variable is given equal weight, then

$$R(f) = \textstyle\sum_{k=1}^{K} |\beta_k| \qquad (3)$$

This approach is particularly useful in the case of high dimensional data in which the selection of regressors that yield the lowest sum of squared residuals within the sample are likely to give some variables that would perform badly in another sample too much influence. Other regulariser functions that are commonly used are the sum of squared parameters (the associated technique is referred to as ridge regression) or a weighted average of the function with absolute values of parameters and the function with squared parameters. The latter is referred to as elastic net regression. Applying these alternative regulariser functions is beyond the scope of the present paper.

An important issue is to set the relative price of variance π. In rigorous LASSO this is done in a way that is grounded in statistical theory and takes into account the possibilities of heteroscedastic, non-Gaussian and cluster-dependent errors (Belloni *et al.*, 2014). Ahrens *et al.* (2020) have introduced a suite of model selection and prediction programs, referred to as *lassopack* in Stata, that include rigorous LASSO and that can be seamlessly compared with estimates obtained by means of classical regression methods.

The results of SML estimation can be found in Table 4. The LASSO coefficients can be directly compared with the corresponding regression coefficients that are also included in the table. In set (1), we pool the data from the two shocks and give all observations equal weight. Hence this situation reduces to finding the best predictors among a set of 146 variables in order to predict 132 social security benefit increase values after an exogenous shock.

Table 4. Supervised Machine Learning Approaches to Identifying Local Predictors of Resilience. Source: Authors' Calculations.

	(1)		(2)		(3)		(4)	
	LASSO; GFC & COVID-19	**Post-estimation OLS GFC & COVID-19**	**Rigorous LASSO; GFC**	**Post-estimation OLS; GFC**	**Rigorous LASSO; COVID-19**	**Post-estimation OLS; COVID-19**	**Rigorous LASSO; GFC & COVID-19**	**Post-estimation OLS; GFC & COVID-19**
Number of observations	132	132	66	66	66	66	132	132
Number of predictors	146	4	15	4	15	4	146	3
ue_r	0.199	0.209***	0.123	0.266***	0.018	0.250***	0.156	0.268***
emp_rate	-0.010	-0.020						
pnoneuro	0.005	0.012**	0.015	0.014	0.004	0.005		
fem_lfpr			-0.019	-0.029				
pubsector_emp	-0.032	-0.057***	-0.019	-0.057***	-0.012	-0.071***	-0.029	-0.076***
pnzdep910					0.007	-0.001		
pdiversity_ind							0.017	0.071***
assuming homoskedasticity	yes	yes	no	No	no	no	no	no
weighting by TA population size	no	no	yes	Yes	yes	yes	yes	yes

Notes: * p<0.05; ** p<0.01; *** p<0.001. Significance levels are based on robust standard errors.

Removal of variables due to perfect collinearity is built in. However, the unconstrained LASSO estimation requires relatively high computational effort (the calculations took about 15 minutes on a high-performance laptop with Intel(R) Core (TM) i7 processor), given that it estimates, besides π, also predictor-specific penalty loadings.

It is clear that LASSO reconfirms the panel data analysis regarding the identification of the most important predictors: they are the census unemployment rate *ue_r* and public sector employment as a share of total employment, *pubsector_emp*. However, two additional variables emerge: *emp_rate* (the percentage of the population aged 15 and over in employment) and *pnoneuro* (the percentage of the population that did not state any European ethnicity among the ethnicities they identified with). As expected, a high value of *emp_rate* reduces benefit uptake while a high value of *pnoneuro* increases it.

Comparing the LASSO results with OLS it is clear that LASSO is a shrinkage estimator: all LASSO coefficients are closer to zero. This highlights the drawback of SML techniques: the resulting model may yield robust predictors, but the influence of the individual variables may not be correctly estimated. The effect of the employment rate is clearly not statistically significant in the case of OLS.

Given that we have panel data, and that the stochastic disturbances of the regression model are likely to be clustered, rigorous LASSO is a more suitable SML technique than ordinary LASSO. Panel (2) shows the results of applying rigorous LASSO to the GFC data. Again, the unemployment rate and public sector employment are the dominant regressors, but *pnoneuro* and the female labour force participation rate, *fem_lfpr*, play also a role and with a negative sign, as expected (more elastic labour supply). The effect of ethnic composition appears more important after the GFC shock than after the COVID-19 shock.

The danger of interpreting individual LASSO coefficients as behavioural parameters is clearly seen by comparing the coefficient of ue_r in panel (3) with all other estimates of this coefficient in this paper (from an average that is greater than 0.2 down to 0.018 in panel (3) of Table 4). This is possibly related to the introduction of the deprivation variable *pnzdep910* into the model: the correlation between the two variables is relatively high (0.83). *pnoneuro* features also in panel (3) but is no longer present when the GFC and COVID-19 data are pooled. This can be seen from panel (4). Here, industrial specialization *pdiversity_ind* returns again as an influential variable. This is probably because the currently available rigorous LASSO procedures have yet to fully encompass parameter estimation of spatial spillover effects (but see Higgins and Martellosio, 2023, for a recent contribution). The spatial panel estimations of the previous section showed that, once spatial spillovers have been accounted for, *pdiversity_ind* is no longer statistically significant.

5. CONCLUSION

In this paper, we focus on the spatial variation across NZ TAs in the initial socio-economic impact – in terms of uptake of social security benefits – of the GFC and the onset of the COVID-19 pandemic. To the best of our knowledge, this is the first study in the international literature that pools regional level data from the GFC and from the pandemic. Using a two-period panel of 66 TAs observed in 2008-09 and 2019-20, we find that, despite the totally different nature of the two shocks, the percentage increase in benefit uptake due to the onset of the COVID-19 pandemic and due to the GFC were similar. This is a statistical coincidence, and any future shock may yield different percentage changes. However, there is a relatively high correlation between the impact of the GFC shock on TAs and the impact of the COVID-19 shock. This persistence in the spatial pattern may continue after a future, as yet unknown shock, that impacts TA labour markets. We link the social security data with 146 indicator variables across 15 domains that are obtained from population censuses that were held about 2 years before each of the two shocks.

To identify urban characteristics that point to economic resilience, we formulate spatial panel regression models guided by stepwise model selection procedures. Additionally, we use ML techniques – given that stepwise regression modelling can lead to over-fitted specifications.

We find that the most resilient TAs had two years previously: (1) a low unemployment rate; and (2) a large public sector. Additionally, but with less predictive power, we find that TAs had a smaller increase in social security uptake after the shock when they had previously: (3) a high employment rate (or high female labour force participation rate); (4) a smaller proportion of the population stating ethnicities other than NZ European; (5) a smaller proportion of the population living in more deprived area units. We also find that interregional spatial spillovers matter and that there are spatial clusters of resilient TAs.

Our results point to a challenge for labour market policy and regional policy, given that regional disparities in unemployment rates are rather persistent. A place-based approach may then be needed to address regional disparities (Van Dijk and Edzes, 2016). The results suggest that greater spreading of public sector employment across the TAs may be helpful in dampening the effect of exogenous shocks. This is consistent with Faggian *et al.* (2018) and Webber *et al.* (2018). The latter found that "regions with greater employment shares in sectors that are less susceptible to demand fluctuations are likely to experience more stable growth rates and be more resilient to economic downturns" (p. 355). Public sector employment shares in TAs varied in our data from 6 per cent to 29 per cent, but these employment shares were virtually constant between the GFC and the COVID-19 pandemic. Clearly, if austerity measures were to be introduced in future years that lead to less public sector employment across all regions, either to reduce public debt or to fund tax cuts, our results do point to a likely decline in regional resilience.

Resilience to exogenous shocks may also be built through active labour market policies (ALMP), which have been shown to be effective in reducing structural unemployment (Miyamoto and Suphaphiphat, 2021; Sahnoun and Abdennadher, 2018; Vooren *et al.*, 2019). This is particularly true at the regional level where tailored ALMP and other place-sensitive policies have enjoyed some success (Wapler *et al.*, 2018).

Hartel *et al.* (2023) note that the vulnerability of a local labour force is likely to depend on the nature of the shock. This is confirmed in the present analysis by there being not only overlaps in predictor variables of the increase in social security uptake between GFC and COVID, but also some differences. Hartel *et al.* (2023) argued that rather than implementing policies to diversify the local economy in a broad sense, a so-called Smart Specialization Strategy (particularly when combined with skill development and enhanced mobility), may well be more effective. This is consistent with our finding that greater industrial diversity yielded less, rather than greater, resistance to the GFC and COVID-19 shocks.

There are many ways in which the present analysis can be extended. The time window considered only covered the initial six months after the shock. Hence the present paper focuses only on the resistance aspect of regional resilience and not the recovery phase. An analysis with a longer time frame could assess the implications of the varying levels of subsequent restrictions on mobility and behaviour, including the reopening of the border. However, such analysis should take endogenous change in predictors of resilience into account.

Finally, it will be important for policy evaluation to move from the regional level of analysis to micro-level local labour market analysis that accounts for heterogeneity across TAs in the impact on employment status, industry, occupation, mobility, etc. of individuals, following an exogenous shock. A recent example is Celbiş et al. (2023) who use machine learning methods to identify individual-level factors that point to vulnerable older-age cohorts in European labour markets after the COVID-19 outbreak. In the NZ context, the required microdata are available through the Integrated Data Infrastructure (IDI) of administrative and survey data, collected and managed by Statistics New Zealand (Stats NZ). It is expected that new developments in machine learning may be particularly helpful in formulating models for identifying determinants of local level resilience by means of very large and complex datasets.

REFERENCES

Ahrens, A., Hansen, C. B. and Schaffer, M. E. (2020). Lassopack: Model selection and prediction with regularized regression in Stata. *The Stata Journal*, 20(1), pp. 176-235.

Andrews, D., Deutscher, N., Hambur, J. and Hansell, D. (2020). The career effects of labour market conditions at entry. OECD Productivity Working Papers, Paris: OECD.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2), pp. 608-650.

Barrett, P. and Poot, J. (2023). Islands, remoteness and effective policy making: Aotearoa New Zealand during the COVID-19 pandemic. *Regional Science Policy and Practice*, 15(3), pp. 682-704.

Bauer, A. and Weber, E. (2021). COVID-19: how much unemployment was caused by the shutdown in Germany? *Applied Economics Letters*, 28(12), pp. 1053-1058.

Blanchflower, D. G. (2004). Self-employment: more may not be better. NBER Working Paper 10286. Cambridge: National Bureau of Economic Research.

Borland, J. (2020). Scarring effects: a review of Australian and international literature. *Australian Journal of Labour Economics*, 23(2), pp. 173-188.

Brada, J. C., Gajewski, P. and Kutan, A. M. (2021). Economic resiliency and recovery, lessons from the financial crisis for the COVID-19 pandemic: A regional perspective from Central and Eastern Europe. *International Review of Financial Analysis*, 74, 101658.

Brown, L. and Greenbaum, R. T. (2017). The role of industrial diversity in economic resilience: An empirical examination across 35 years. *Urban Studies*, 54(6), pp. 1347-1366.

Celbiş, M. G., Wong, P. H., Kourtit, K. and Nijkamp, P. (2023). Impacts of the COVID-19 outbreak on older-age cohorts in European labor markets: A machine learning exploration of vulnerable groups. *Regional Science Policy and Practice*, 15(3), pp. 559-584.

Cho, S. J., Lee, J. Y. and Winters, J. V. (2021). Employment impacts of the COVID-19 pandemic across metropolitan status and size. *Growth and Change*, 52(4), pp. 1958-1996.

Clyne, D. V. and Smith, T. G. (2022). Economic insecurity during the COVID-19 pandemic: insights from the Great Recession. *New Zealand Economic Papers*, 56(1), pp. 81-89.

Cochrane, W. and Poot, J. (2009). Regional labour market adjustment and social security benefit uptake. Proceedings of the Joint LEW13/ALMRW Conference, Wellington, pp. 343-354.

Cochrane, W. and Poot, J. (2020). Did the post-1986 decline in the homeownership rate benefit the New Zealand labour market? A spatial-econometric exploration. *Asia-Pacific Journal of Regional Science*, 4, pp. 261-284.

Dijkstra, L., Garcilazo, E. and McCann, P. (2015). The effects of the global financial crisis on European regions and cities. *Journal of Economic Geography*, 15(5), pp. 935-949.

Dyason, D., Fieger, P. and Rossouw, R. (2021). Covid-19, the effect of lockdowns on retail expenditure and displacement effects on the regional economy. *Australasian Journal of Regional Studies*, 27(1), pp. 66-87.

Elhorst, J. P. (2014). *Spatial Econometrics: From Cross-sectional Data to Spatial Panels*, Springer, Heidelberg.

Faggian, A., Gemmiti, R., Jacquet, T. and Santini, I. (2018). Regional economic resilience: the experience of the Italian local labor systems. *Annals of Regional Science*, 60, pp. 393-410.

Fletcher, M., Prickett, K.C., and Chapple, S. (2022). Immediate employment and income impacts of Covid-19 in New Zealand: evidence from a survey conducted during the Alert Level 4 lockdown. *New Zealand Economic Papers*, 56(1), pp. 73-80.

Furnival, G. M. and Wilson, R. W. (1974). Regression by leaps and bounds. *Technometrics*, 16, pp. 499–511.

Gauld, R. (2023). A review of public policies on Covid-19: the New Zealand experience. *Public Administration and Policy: An Asia-Pacific Journal*, 26(1), pp. 10-20.

Giannakis, E., and Bruggeman, A. (2017). Determinants of regional resilience to economic crisis: a European perspective. *European Planning Studies*, 25(8), pp. 1394-1415.

Groenewold, N. (2020). The response of the Australian states to a national economic shock: a statistical analysis of regional economic resilience. *Australasian Journal of Regional Studies*, 26(2), pp. 186-208.

Halleck Vega, S. and Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, 55(3), pp. 339-363.

Hamann, S., Niebuhr, A., Roth, D. and Sieglen, G. (2023). How does the COVID-19 pandemic affect regional labour markets and why do large cities suffer most? IAB-Discussion Paper, No. 3/2023, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

Hartal, S., Kourtit, K. and Pascariu, G. C. (2023). Spatial labour markets and resilience of cities and regions: a critical review. *Regional Studies,* 57(12), 2359-2372.

Hassink, R. (2010). Regional resilience: a promising concept to explain differences in regional economic adaptability? *Cambridge Journal of Regions, Economy and Society*, 3(1), pp. 45–58.

Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*, Boca Raton, CRC Press.

Hershbein, B. J. and Stuart, B.A. (2020). The enduring local harm from recessions. *W.E. Upjohn Institute - Employment Research Newsletter*, 27, pp. 1-4.

Higgins, A. and Martellosio, F. (2023). Shrinkage estimation of network spillovers with factor structured errors. *Journal of Econometrics*, 233(1), pp. 66-87.

Houston, D. (2020). Local resistance to rising unemployment in the context of the COVID-19 mitigation policies across Great Britain. *Regional Science Policy and Practice*, 12, pp. 1189-1209.

Hundt, C. and Grün, L. (2022). Resilience and specialization – How German regions weathered the Great Recession. *ZFW – Advances in Economic Geography*, 66, pp. 96-110.

Hyslop, D. R., Maré, D. C. and Minehan, S. (2023). COVID-19 wage subsidy: outcome evaluation. Motu Working Paper 23-03. Motu Economic and Public Policy Research, Wellington.

Kim, A., Lim, J. and Colletta, A. (2023). How regional economic structure matters in the era of COVID-19: resilience capacity of U.S. states. *Annals of Regional Science*, 70, pp. 159-185.

Klein, E., Cook, K., Maury, S. and Bowey, K. (2021). Gendered impacts of changing social security payments during COVID-19 lockdowns: an exploratory study. *Australian Journal of Labour Economics*, 24(2), pp. 213-225.

Lindsey, C. and Sheather, S. (2010). Variable selection in linear regression. *The Stata Journal*, 10(4), pp. 650-669.

Markham, S. (2020). Monetary policy and regional unemployment. Analytical Notes AN2020/8. Reserve Bank of New Zealand.

Martin, R. and Sunley, P. (2015). On the notion of regional economic resilience: conceptualization and explanation. *Journal of Economic Geography*, 15(1), pp. 1-42.

McNeish, D. M. (2015). Using Lasso for predictor selection and to assuage overfitting: a method long overlooked in behavioural sciences. *Multivariate Behavioral Research*, 50(5), pp. 471-484.

Miyamoto, H. and Suphaphiphat, N. (2021). Mitigating long-term unemployment in Europe. *IZA Journal of Labor Policy*, 11(1), 1-27.
Modica, M., and Reggiani, A. (2015). Spatial economic resilience: overview and perspectives. *Networks and Spatial Economics*, 15, pp. 211-233.
Molina, M. and Garip, F. (2019). Machine learning for sociology. *Annual Review of Sociology*, 45, pp. 27-45.
Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), pp. 87-106.
Plater, V. and Claridge, M. (2000). Facts about economic integration: how integrated is New Zealand with the rest of the world? New Zealand Treasury Working Paper No. 00/21. The Treasury, Wellington.
Rea, D. and Maloney, T. (2021). Nowcasting the current rate of unemployment using administrative data. AUT Centre for Social Data Analytics.
Rota, F. S., Bagliani, M. and Feletig, P. (2020). Breaking the black-box of regional resilience: a taxonomy using a dynamic cumulative shift-share occupational approach. *Sustainability*, 12(9070), 1-27.
Salmond, A., Crampton, P. and Sutton, F. (1998). NZDep91: A New Zealand index of deprivation. *Australian and New Zealand Journal of Public Health*, 22(7), pp. 835-837.
Sahnoun, M. and Abdennadher, C. (2018). The assessment of active labor market policies: evidence from OECD countries. *Economia Politica*, 35, pp. 257-283.
Statistic New Zealand (2022). Investigating the relationship between unemployment and benefit receipt. Statistics New Zealand.
Van Dijk, J. and Edzes, A. (2016). Towards inclusive and resilient regional labour markets: challenges for research and policy. *Investigaciones Regionales – Journal of Regional Research*, 36, pp. 169-190.
Vooren, M., Haelermans, C., Groot, W. and Maassen van den Brink, H. (2019). The effectiveness of active labor market policies: A meta-analysis. *Journal of Economic Surveys*, 33(1), pp. 125-149.
Wapler, R., Werner, D. and Wolf, K. (2018). Active labour market policies in Germany: do regional labour markets benefit? *Applied Economics*, 50(51), pp. 5561-5578.
Webber, D. J., Healy, A. and Bristow, G. (2018). Regional growth paths and resilience: A European analysis. *Economic Geography*, 94(4), pp. 355-375.